CHILD DEVELOPMENT

Vol. 27
June, 1956
No. 2

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS

OF THE

SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.

Contents for June

Editorial Note
William E. Martin 83

The Stimulus Generalization Gradient as a Function of the Intensity of Stimulus Lights
Charles C. Spiker 85

The Effects of Differential Reinforcement and Motivation on Prediction Responses of Children
Sidney Rosenblum 99

An Objective Method for the Measurement and Analysis of Child-Adult Interaction
Clark E. Moustakas, Irving E. Sigel and Henry D. Schalock 109

Identification with Parents as a Determinant of Doll Play Aggression
Harry Levin and Robert R. Sears 135

The Development of Prejudice: A Critique of the *Horowitz Faces Test*
Eugene E. Levitt 155

The Effect of Concept Knowledge on Discrimination Learning
Winifred O. Shepard and Maurice Schaefer 173

SUPPLEMENT: CHICAGO CONFERENCE ON CHILD DEVELOPMENT, 1955 17

Information for Contributors and Subscribers

Child Development, one of three publications issued by the Society for Researc in Child Development, Inc., is issued quarterly in March, June, September an December. The subscription price is $6.00 per year, postpaid.

Child Development Abstracts and Bibliography is now in its thirtieth volum This publication is issued February through December—six numbers to a volum The subscription price is $5.00 per year, postpaid.

Monographs of the Society for Research in Child Development, Inc., issue irregularly during the year, consists primarily of detailed critical reviews and seri atim studies. The subscription price is $4.00 per yearly volume.

The three publications of the Society are available at the special price of $12.0 per year.

A limited number of back issues of all publications is available at reduced price A mimeographed list may be had upon request.

Contributors are asked to refer to the *Publication Manual of the American Psy chological Association* (Supplement to the *Psychological Bulletin*, 1952, 49, No. Part 2) and *A Manual of Style* (University of Chicago Press, 1950) in the preparatio of manuscripts.

Manuscripts and *all business correspondence* relating to the publications shoul be addressed to Child Development Publications, Purdue University, Lafayett Indiana.

CHILD DEVELOPMENT

Vol. 27
June, 1956
No. 2

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

REGINALD MACGREGOR ARCHIBALD
Rockefeller Institute for Medical Research

JOSEPH H. BRITTON
Pennsylvania State University

SIDNEY W. BIJOU
University of Washington

VAUGHN J. CRANDALL
Fels Research Institute

ERCEL S. EPPRIGHT
Iowa State College

C. GLEN KING
Nutrition Foundation

BENJAMIN PASAMANICK
Ohio State University

ANSELM STRAUSS
University of Chicago

RALPH L. WITHERSPOON
Florida State University

Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Printed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

EDITORIAL NOTE

This issue of *Child Development* is double-sized, since it includes a Supplement on the Chicago Conference on Child Development held on October 26-29, 1955, under the joint sponsorship of the Elizabeth McCormick Memorial Fund and the Society for Research in Child Development with Donald Brieland, Director of the Fund, and William W. Greulich, President of the Society, as co-chairmen.

The stated purposes of the Conference were: (*a*) to give an opportunity for participating institutions to summarize their program interests and research-in-progress; (*b*) to indicate data which they have available which may be of interest to other people in the field of child development; (*c*) to explore the feasibility of pooling certain data for research purposes; (*d*) to exchange research ideas of the people represented, including consideration of problems in which they are interested but with which they do not plan to deal themselves.

Prior to the Conference, representatives of each participating university or research group prepared and circulated to all participants written summary descriptions of research-in-progress, data available and publications. Summaries were prepared by the following:

Nancy Bayley, National Institute of Mental Health, National Institutes of Health, Public Health Service, U.S. Department of Health, Education and Welfare

Dale B. Harris, Institute of Child Welfare, University of Minnesota

Harold E. Jones, Institute of Child Welfare, University of California

Boyd R. McCandless, Iowa Child Welfare Research Station, State University of Iowa

Jean Walker Macfarlane, Guidance Study, Institute of Child Welfare, University of California

Lester W. Sontag, Fels Research Institute, Antioch College

John W. M. Whiting, Laboratory of Human Development, Harvard University

These summaries provided the basis for three discussion sessions on Hypotheses, Methodology and Implications for Theory in Current Research which were led respectively by John D. Benjamin, Denver Child Research Council; Roger G. Barker, Department of Psychology, University of Kansas; and Alfred L. Baldwin, Department of Child Development and Family

Relationships, Cornell University. These papers and the ensuing discussions do not appear in the Supplement.

A second set of papers dealing with the research concerns and activities of individuals was also prepared and distributed prior to the conference. These papers, together with two delivered at the Conference by John E. Anderson and Harold E. Jones, constitute the Supplement. They have been revised and abridged to meet space requirements. A longer paper prepared for the Conference by Helen L. Koch, Department of Psychology, University of Chicago, will appear in a later issue of *Child Development.*

The Elizabeth McCormick Memorial Fund underwrote the costs of the Conference and made possible the publication of the Supplement. To the Fund and its Director go the gratitude of the members of the Society for Research in Child Development and the readers of *Child Development.*

WILLIAM E. MARTIN
Editor

THE STIMULUS GENERALIZATION GRADIENT AS A FUNCTION OF THE INTENSITY OF STIMULUS LIGHTS

CHARLES C. SPIKER[1]
Iowa Child Welfare Research Station

The experiment reported here is the second in a series designed to develop procedures and apparatus appropriate for the study of the generalization behavior of children. In the first experiment (10), the amount of generalization was shown to be an increasing function of the number of reinforcements given on the training stimulus. The present experiment is concerned with the hypothesis that the shape of the generalization gradient, for stimulus intensity dimensions, is dependent upon the absolute values of the stimuli selected as well as upon the degree to which the stimuli are separated on the stimulus dimension.

Several investigators have found that if the subject (*S*) is conditioned to an intense stimulus and is then tested on stimuli of lesser intensity, the generalization gradient is steeper than if the *S* is conditioned to a weak stimulus and tested on more intense stimuli. Hovland (3) reported this finding with the GSR for human *S*s using auditory stimuli; Brown (1) and Perkins (8) obtained similar findings for an instrumental approach response of rats to lights differing in brightness; and Grice and Saltz (2) demonstrated the phenomenon in rats' approach responses to stimuli differing in size. Hull (4) attempted to explain some of these results in terms of his principle of stimulus dynamism. Although the details of his explanation are not given here, it involves the assumption that the intensity of a stimulus contributes to the general motivational level of *S*, and that the motivational factors multiply with the generalized habits. Thus, given two stimuli (S_1 and S_2) that are equally separated from the conditioned stimulus, S_1 being more intense and S_2 less intense than the conditioned

[1] The writer is indebted to Miss Kathryn Norcross and Mr. Lewis P. Lipsitt for aid in collecting the data and for valuable suggestions regarding procedural matters.

stimulus, S_1 will have a greater reaction potential than will S_2, The deductive consequence of these assumptions, and others in the system, is that the generalization gradient will be steeper when the test is made with less intense stimuli than when made with more intense stimuli.

In this experiment, the tendency to respond to the conditioned stimulus was established by giving differential reinforcement during training, in which *S* was reinforced for responding to the conditioned stimulus and was nonreinforced for responding to another stimulus that was different from the conditioned stimulus in terms of one characteristic (hue), but the same with respect to the dimension being investigated (brightness). Subsequently, *S*'s tendency to respond to other stimuli, differing from the training stimulus only in brightness, was determined in a generalization test. Two levels of reinforcement during training were also used in an attempt to confirm the findings of the previous experiment (10).

Method

Subjects and Experimental Design

The *S*s were 60 children from the Iowa Child Welfare Research Station preschool laboratories. They ranged in age from 3 years, 9 months, to 5 years, 8 months. Most *S*s had been used in a previous study of generalization (10) in which the present apparatus was used, but in which the stimuli differed in hue rather than in brightness.

The experimental design was such that one-half the *S*s were randomly assigned to a group that was trained to respond to a bright light and then tested on dimmer lights. The other half were assigned to a group trained to respond to a dim light and tested on brighter lights. One-half of each of these groups was randomly assigned to a group that received 12 presentations of the postive (white) light and 12 presentations of the negative (blue) light during reinforcement training. The other halves of the two groups received 24 presentations of each of the positive and negative stimuli. For convenience, these four subgroups will be referred to as B-12 (bright light, 12 presentations of each stimulus); B-24 (bright, 24 presentations); D-12 (dim, 12 presentations); and D-24 (dim, 24 presentations).

Apparatus

The experiment was conducted in a semi-darkened experimental room with an adjoining observation room. The apparatus, which has been described elsewhere in detail (10), was placed in the window between the two rooms. *S* operated the apparatus from the experimental room; the experimenter (*E*) operated the controls and recorded *S*'s responses from the observation booth. S faced a gray metal panel, 20 in. wide and 12 in. high; 3 in. from the top was centered an aperture, 2 in. in diameter, covered with flashed opal glass. Below and to the left of the aperture was a transparent

container into which marbles were ejected. A metal rod with a hand grip extended from the panel below and to the right of the aperture. Directly below the apparatus was a wooden stand, 2 ft. square and 8 in. high, on which *S* stood during the experiment.

The apparatus was enclosed in a metal cabinet, 12 in. high, 20 in. wide, and 12 in. deep. Approximately 7 in. behind the aperture were six 75-watt projection lamps, arranged so that their beams focused on the aperture. Each lamp was enclosed in a reflecting container on the front of which 2 in. square wooden diaphragms or glass color filters could be mounted. A stepping relay controlled the current to the lamps and the relay was in turn controlled by two interval timers. One of the timers controlled the presentation interval and the other controlled the interval between presentations. A multi-contact selector switch was used to determine which stimuli would be presented and the order of their presentation.

If a stimulus lamp was on at the time the response occurred, the response was recorded on one of six electrical impulse counters. Each counter was associated with one stimulus, so that the responses to each stimulus were separately recorded. Approximately 8 pounds of pressure were required to move the lever the ¾ in. necessary to activate the counters and the marble ejector. The marble ejector was activated by a response only on those trials during which the reinforced stimulus was presented.

Stimuli

The stimuli were lights, different brightnesses being obtained by placing wooden diaphragms between the lamps and the stimulus aperture. The brightnesses of four stimuli, measured at the aperture with a McBeth Illuminometer, were approximately 2.5, 25, 80, and 250 ft.-candles. The difference between the common logarithms of 2.5 and 25, and of 25 and 250 is 1.0. The 80 ft.-candle stimulus is one-half a logarithm unit between the 25 and 250 ft.-candle stimuli. One-half the *S*s were trained with the 250 stimulus, while the other half were trained with the 2.5 stimulus. Two additional stimuli were constructed by placing a blue glass color filter over the diaphragms for the 2.5 and 250 ft.-candle stimuli. This produced two pairs of stimuli, the members of each pair approximately equal in brightness but differing in hue. The blue stimuli were used as the nonreinforced (negative) stimuli in the training sessions.

Preliminary Training.

S was brought to the experimental room, shown a variety of dime-store toys that were spread out on a small table, and was allowed to choose the toy that he wished to have. *E* informed him that he must buy the toy with marbles and that *E* would show him how to get the necessary marbles. The apparatus was turned on and set to present the training stimulus series. This series consisted of an equal number of separate presentations of the

white (positive) and of the blue (negative) stimuli. Groups B-12 and B-24 were presented with a white and blue light of approximately 250 ft.-candles; Groups D-12 and D-24 received a white and a blue light of approximately 2.5 ft.-candles. Throughout the experiment, the presentation period was 3 seconds and the inter-presentation interval was 2½ seconds.

E first demonstrated the apparatus by pulling the lever several times while the training light was on and pointing to the container into which the marbles were being ejected. When the blue light came on, *E* pulled the lever a few times to demonstrate that no marbles were being ejected. *S* was then allowed to pull the lever to obtain marbles. The preliminary training was continued until *S* learned to pull more than once when the training light was on and failed to pull the lever between stimulus presentations. As soon as *S* had met these criteria, *E* stopped the apparatus and returned the stimulus presentation device to the starting position in preparation for the reinforcement training.

Reinforcement Training

E returned *S* to his position before the apparatus, and informed him that *E* was going into the next room to start the apparatus again and that *S* would be informed when he had obtained enough marbles to buy the toy. *E* went into the observation booth and started the apparatus. The reinforcement stimulus series was the same one given during the preliminary training. In each successive block of six presentations, there were three each of the positive and negative stimuli. Neither stimulus was presented more than twice in succession. Groups B-12 and D-12 were given 12 presentations of each the white and the blue lights and Groups B-24 and D-24 were given 24 presentations of each. *E* recorded on prepared sheets the number of responses *S* made to each stimulus presentation.

Generalization Test

Following the last trial on the reinforcement training, *E* turned the selector switch so that the generalization stimulus series was presented. The time between the last reinforcement training trial and the first generalization test trial was eight seconds. The generalization series was introduced with no other changes in the procedure.

The series consisted of two cycles, each cycle having seven presentations of the positive stimulus with each response being reinforced, and five non-reinforced presentations of the test stimuli. The test stimuli were presented on trials 2, 4, 7, 9, and 12 of the cycle. The test stimuli included the positive stimulus, the negative stimulus, and each of the other three stimuli differing from the positive stimulus in brightness. The test stimuli were presented in counterbalanced order so that within each group each test stimulus appeared in each position an equal number of times. For each *S*, the second cycle was the same as his first.

Results

Responses to the Reinforced Stimulus during the Generalization Series

An analysis was first made of the number of responses to the reinforced training stimulus during the generalization test. This analysis should reveal the differential effects on learning, if any, of the different conditions that prevailed during the acquisition stages. A Lindquist Type III analysis (6, p. 281) was conducted, in which the main effects were the number of reinforced trials during training, the brightness of the conditioned stimulus, and successive three-trial blocks for the first 12 of the 14 reinforced trials of the generalization test. The results of the analysis need not be presented in detail, since only the trial-block effect proved significant at less than the 5 per cent level. This significant effect reflects the fact that the number of responses to the reinforced conditioned stimulus tended to decrease with an increase in the number of previous test trials. This decrease is probably due to the generalized effects of nonreinforcement of the test stimuli.

This analysis makes tenable the hypothesis that the four groups did not differ significantly in performance to the conditioned stimulus as a function of the different conditions that obtained during training. This appears to be true both for the brightness of the conditioned stimulus and for the number of reinforcements received during training.

Responses to the Test Stimuli for the Complete Generalization Test

Table 1 gives the mean number of responses made to each test stimulus for the first and second cycles of the generalization test series. An examina-

Table 1

MEAN NUMBER OF RESPONSES TO EACH TEST STIMULUS

Test Stimulus	First Cycle Training Reinforcements			Second Cycle Training Reinforcements		
	12	24	Mean	12	24	Mean
BRIGHT-TRAINED GROUPS						
250	2.40	3.33	2.87	2.00	2.40	2.20
80	1.13	1.60	1.37	1.13	0.93	1.03
25	0.80	1.13	0.97	1.00	0.67	0.83
2.5	0.80	0.53	0.67	0.33	0.53	0.43
Mean	1.28	1.65	1.47	1.12	1.13	1.12
DIM-TRAINED GROUPS						
2.5	2.33	1.80	2.07	1.93	2.27	2.10
25	1.40	1.87	1.63	1.60	1.07	1.33
80	1.40	1.27	1.33	0.87	1.13	1.00
250	1.07	1.47	1.27	0.67	1.00	0.83
Mean	1.55	1.60	1.58	1.27	1.37	1.32

tion of this table reveals (*a*) a somewhat greater mean number of responses for the Groups B-24 and D-24 than for D-12 and B-12; (*b*) a greater mean number of responses for all groups combined during the first cycle than during the second cycle; (*c*) a slightly greater mean number of responses to the test stimuli for Groups D-24 and D-12 than for Groups B-24 and B-12; (*d*) a decrease, for all groups combined, in the number of responses to the test stimuli as the difference between training and test stimuli increases; and (*e*) a relatively greater decrease in these responses for Groups B-12 and B-24 than for Groups D-12 and D-24. The appropriate analysis for testing the statistical significance of these observations is an extension of Lindquist's Type VI design (6, p. 292). The main factors in the analysis are the number of reinforcements during training, the brightness of the training stimulus, the different test stimuli, and the cycles of the generalization test.

In order to analyze the data, the stimulus dimension was reversed for *Ss* trained to the dim stimulus. That is, the responses made by these *Ss* to the 2.5 ft.-candle stimulus, their training stimulus, were compared with the

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF RESPONSES TO EACH TEST STIMULUS PRESENTATION

Source	*df*	*Mean Square*	*F*	*p*
Subjects	59			
Reinforcement (R)	1	1.88		
Training Stimulus (T)	1	4.01		
R × T	1	0.40		
error (b)	56	4.14		
Within Subjects	300			
Test Stimuli (S)	2	73.51	41.53	<.001
Cycles (C)	1	7.51	11.73	<.005
S × C	2	0.12		
S × T	2	9.88	5.58	<.005
S × R	2	0.68		
C × T	1	0.28		
C × R	1	0.54		
error (w)	280			
$error_1$ (w)	56	0.64		
$error_2$ (w)	112	1.77		
$error_3$ (w)	112	1.06		
Total	359			

NOTE.—Higher order interactions have been omitted, since none was significant at less than the 10 per cent level. The stimulus dimension was reversed for *Ss* trained to the dim stimulus, and responses of all *Ss* to the 80 ft.-candle stimulus have been omitted.

responses made by the bright-trained *Ss* to their 250 ft.-candle training stimulus. Responses made by the dim-trained group to the 25 ft.-candle stimulus were compared with those made by the bright-trained group to the same stimulus; and responses made by the dim-trained group to the 250 ft.-candle stimulus were compared with those made by the bright-trained group to the 2.5 ft.-candle stimulus. By omitting the responses of both groups to the 80 ft.-candle stimulus, there are equal logarithmic steps between the training and test stimuli for both groups.

A summary of the analysis of variance is presented in Table 2. The triple and quadruple interactions have been omitted from the table since

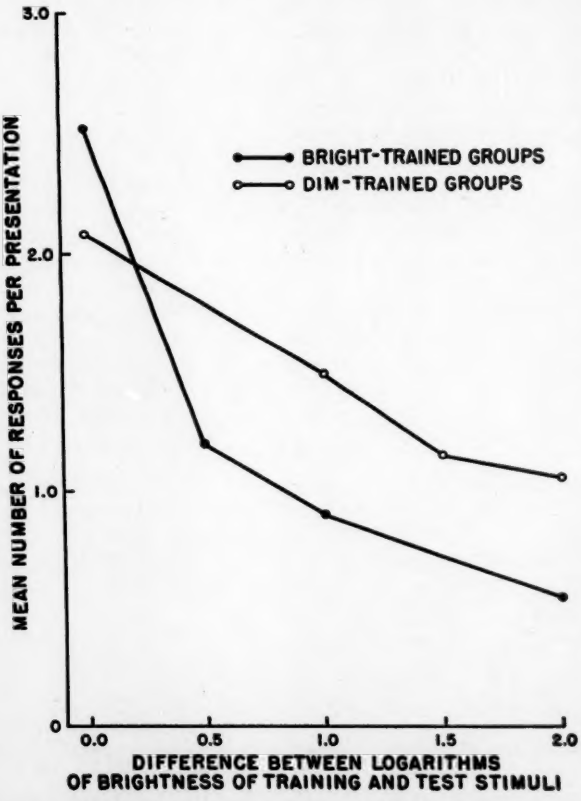

FIGURE 1—Mean number of responses per test stimulus for complete generalization series.

none was significant at less than the 10 per cent level. As the table indicates, neither the reinforcement factor (R) nor the brightness of the training stimulus (T) proved to be significant. The significant cycle effect indicates the reduction in the number of responses made to the test stimuli from the first to the second cycle. The significant test stimulus effect indicates that the generalization gradient for the combined groups cannot be attributed to chance. The significant S × T interaction reflects the fact that the generalization gradient for *Ss* trained to the bright stimulus is significantly steeper than the gradient for *Ss* trained to the dim stimulus. This latter finding is illustrated in Figure 1, where the means of the numbers of responses per test stimulus presentation, for the pooled bright-trained and pooled dim-trained groups, are plotted against the difference between the common logarithms of the training and test stimuli. The mean number of responses to the 80 ft.-candle stimulus is also included for each group. It is apparent from Figure 1 that the gradient for the bright-trained *Ss* is steeper than that for the dim-trained.

In order to establish the existence of a generalization gradient for the bright-trained and the dim-trained groups, Lindquist's Type I analysis (6, p. 267) was made separately for these two groups. In this analysis, responses to the 80 ft.-candle stimulus were included. The response measure was the mean number of responses to each test stimulus for the entire generalization test series (disregarding cycles) and the level of reinforcement during training was ignored. Table 3 presents a summary of these analyses. The effect of primary interest here is that for test stimuli, since a significant test stimulus effect indicates a significant generalization gradient. For both groups, the generalization gradient proves to be highly significant ($p < .001$). Table 4 gives the probabilities resulting from t tests of the significance

TABLE 3

SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF RESPONSES TO THE TEST STIMULI FOR THE BRIGHT-TRAINED AND THE DIM-TRAINED GROUPS SEPARATELY

Source	*df*	*Mean Square*	*F*	*p*
	BRIGHT-TRAINED GROUPS			
Test Stimuli (S)	3	98.48	40.53	<.001
Subjects (I)	29	9.46		
S × I	87	2.43		
	DIM-TRAINED GROUPS			
Test Stimuli (S)	3	25.69	7.56	<.001
Subjects (I)	29	12.02		
S × I	87	3.40		

NOTE.—This analysis includes the responses to the 80 ft.-candle stimulus.

of the differences between the means for pairs of stimuli for the two groups separately. It appears that the significance of the gradient is primarily due to differences between responses to the training stimulus (under test conditions) and the other test stimuli. For the bright-trained group, however, there is a significant difference ($p < .005$) between the means for the 80 ft.-candle and the 2.5 ft.-candle stimuli.

TABLE 4

PROBABILITIES OBTAINED FROM *t* TESTS COMPARING THE MEANS FOR PAIRS OF TEST STIMULI FOR THE BRIGHT-TRAINED AND DIM-TRAINED GROUPS SEPARATELY

Bright-Trained Group		Dim-Trained Group	
Stimulus Pairs	*p*	*Stimulus Pairs*	*p*
250 vs. 80	<.001	2.5 vs. 25	<.02
250 vs. 25	<.001	2.5 vs. 80	<.001
250 vs. 2.5	<.001	2.5 vs. 250	<.001
80 vs. 25	<.20	25 vs. 80	<.20
80 vs. 2.5	<.005	25 vs. 250	<.10
25 vs. 2.5	<.10	80 vs. 250	<.70

Responses to the Initial Test Stimulus Presentation

The failure to find a significant effect for reinforcement during training, contrary to the results of other studies, suggested that the additional reinforcements given during the generalization test may have obliterated any differences that existed initially between the two reinforcement groups. Accordingly, the number of responses made to the first test stimulus presented was obtained for each *S*. There were three *S*s in each of the four main groups that received the same test stimulus on the first presentation. Since one of these subgroups received the negative stimulus first and another received the 80 ft.-candle stimulus, neither of which are of concern here, there remained 36 *S*s, nine in each group, who could be used in this analysis. The means of these cases, each cell based on three *S*s, are presented in Table 5. Trends corresponding to those in Table 1 may be observed here. A Lindquist Type III analysis (6, p. 281) was made of these data and the summary is presented in Table 6. The test stimulus variable is again highly significant. The data show a significant ($p < .025$) reinforcement effect, and a significant ($p < .05$) interaction between reinforcements and test stimuli. The interaction between brightness of the training stimulus and the test stimuli is not as pronounced as in the first analysis, although the pattern of the corresponding means is the same. In Figure 2, the means for *S*s receiving 12 reinforcements and the means for those receiving 24

reinforcements are plotted as a function of the test stimuli. It is apparent that the reinforcement effect is due to the greater number of generalized responses by the 24 reinforcement *S*s and that the interaction is due to the steeper slope of the gradient for these *S*s. A *t* test of the simple effects shows that the group receiving 24 reinforcements made significantly more responses to the first and second test stimuli ($p < .05$). This analysis thus indicates that the greater number of reinforcements during the training resulted in significantly greater generalization to the stimuli near to the training stimulus, at least during the early stages of the generalization test.

TABLE 5

MEAN NUMBER OF RESPONSES TO FIRST TEST STIMULUS PRESENTED THE SUBJECTS

		Reinforcements During Training		
Training Stimulus	*Test Stimulus*	*12*	*24*	*Mean*
250	250	1.33	3.00	2.17
	25	0.33	1.33	0.83
	2.5	0.67	0.33	0.53
	Mean	0.78	1.56	1.17
2.5	2.5	1.67	2.33	2.00
	25	1.33	2.33	1.83
	250	0.67	0.33	0.50
	Mean	1.22	1.67	1.44

TABLE 6

ANALYSIS OF VARIANCE OF THE MEAN NUMBER OF RESPONSES TO FIRST TEST STIMULUS PRESENTED TO SUBJECTS

Source	*df*	*Mean Square*	*F*	*p*
Training Stimulus (T)	1	0.70		
Reinforcements (R)	1	3.36	5.89	<.025
Test Stimuli (S)	2	7.53	13.21	<.001
T × R	1	0.25		
T × S	2	1.19	2.09	<.20
R × S	2	2.02	3.54	<.05
T × R × S	2	0.25		
Within Cells	24	0.57		
Total	35			

NOTE.—The stimulus dimension has been reversed for *S*s trained to the dim stimulus, and *S*s given the 80 ft.-candle stimulus first have been omitted.

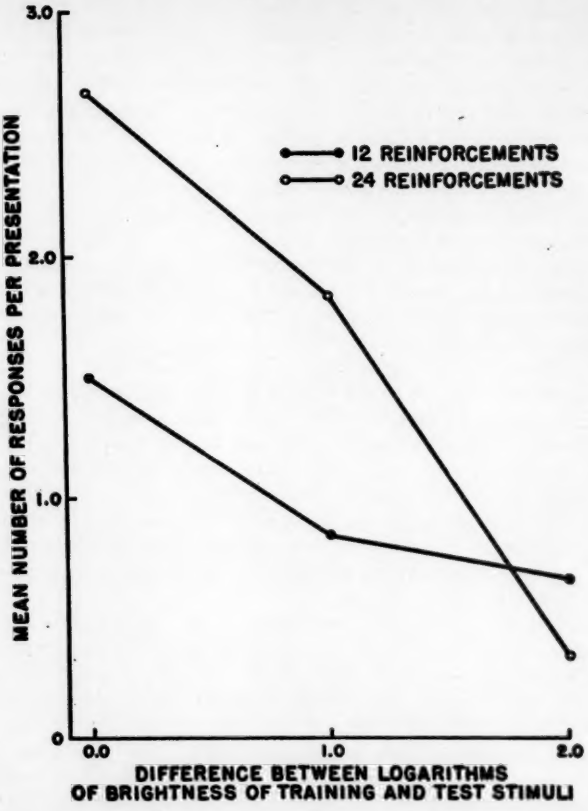

FIGURE 2—Mean number of responses to first test stimulus presentation.

DISCUSSION

Analysis of the number of responses made to the test stimuli during the generalization test reveals a highly significant stimulus generalization gradient for all *Ss* combined. That this "group gradient" is not merely an artifact of combining individual step-wise gradients is demonstrated by the fact that 80 per cent of all *Ss* responded fewer times to the most distant test stimulus than to the nonreinforced training stimulus and 60 per cent responded fewer times to the most distant test stimulus than they did to the test stimulus nearest the training stimulus.

Analysis also indicated that the gradient for *Ss* trained to the bright stimulus is significantly steeper than that for *Ss* trained to the dim stimulus. This finding is in agreement with that of Grice and Saltz (2) for the size dimension, with that of Brown (1) and Perkins (8) for the brightness dimension, and with Hovland's (3) for loudness. It is also in general agreement with the prediction made by Hull (4) on the basis of the principle of stimulus dynamism. Further analysis revealed that the generalization gradient for the dim-trained *Ss* is statistically significant, indicating that a generalization gradient can be demonstrated even when stimulus intensity operates in the opposite direction.

The significant effect of cycles—representing repeated presentations of nonreinforced test stimuli—is also one of the generalization phenomena frequently observed. This principle is one of the main axioms in stimulus-response theories that are designed to account for the various phenomena in discrimination learning (5, 9).

Although the effects of the number of reinforcements during training failed to achieve significance in the above described analysis, a significant reinforcement effect was obtained when only the responses to the first stimulus presentation were analyzed. In addition, a significant interaction between test stimuli and number of reinforcements was found, indicating a steeper gradient for the *Ss* receiving 24 reinforcements than for the group receiving only 12. This result may also be derived from Hull's formulations, and has been empirically demonstrated for the instrumental approach response of rats (7).

The results of this experiment indicate that the apparatus and procedures used offer a fruitful and practical means of investigating generalization phenomena with child *Ss*. The method has been tested on hypotheses derived from the results obtained with more conventional methods and subjects. The hypotheses that were verified deal with the establishment of a gradient of stimulus generalization; the effects of reinforcement on the amount and shape of this gradient; the effects of repeated nonreinforced stimulus presentations on the magnitude of generalization; and the effect of stimulus intensity on the shape of the gradient. It now appears that the method warrants use in the study of some of the generalization phenomena which may best be studied with children; e.g., the effect on the shape of the gradient of prior name-learning experience with the stimuli; the effects on generalization behavior of previous training to generalize or to discriminate; and the relationship between level of maturity (i.e., chronological or mental age) and the magnitude, range, and steepness of the generalization gradient.

Summary

This experiment is primarily concerned with the study of the effects on generalization of the number of reinforcements with the conditioned stim-

ulus and with the effects of the absolute brightness of the stimuli on the magnitude and shape of the generalization gradient. Sixty preschool *Ss* were given differential reinforcement to a white (positive) and a blue (negative) light. For one-half the *Ss*, the positive stimulus was the brightest of four stimuli differing from each other in brightness; for the other half, it was the dimmest. One-half of each these groups was given 12 presentations and the other half was given 24 presentations of each the positive and negative stimuli. The response was the repeated pulling of a lever for marbles during a 3-second presentation of the conditioned stimulus, and the response measure was the number of such responses that occurred during the 3-second periods. Immediately following training, the *Ss* were tested, without reinforcement, on each of the four stimuli differing in brightness. The main findings were (*a*) a statistically significant generalization gradient for all groups combined; (*b*) a steeper gradient for the bright-trained than for the dim-trained *Ss*, and (*c*) a greater number of generalized responses for *Ss* given 24 reinforced training trials than for *Ss* given only 12, and a steeper gradient for the former than for the latter.

References

1. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, 33, 209-226.

2. Grice, R. G., & Saltz, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, 40, 702-708.

3. Hovland, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensity of tone. *J. genet. Psychol.*, 1937, 51, 279-291.

4. Hull, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychol. Rev.*, 1949, 56, 67-76.

5. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

6. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton-Mifflin, 1953.

7. Margolius, G. Stimulus generalization of an instrumental response as a function of the number of reinforced trials. *J. exp. Psychol.*, 1955, 49, 105-111.

8. Perkins, C. C., Jr. The relationship between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, 46, 225-231.

9. SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

10. SPIKER, C. C. The effects of number of reinforcements on the strength of a generalized instrumental response. *Child Develpm.*, 1956, 27, 37-44.

THE EFFECTS OF DIFFERENTIAL REINFORCEMENT AND MOTIVATION ON PREDICTION RESPONSES OF CHILDREN[1]

SIDNEY ROSENBLUM[2]
Iowa Child Welfare Research Station

Since the early 1930's the effects on behavior of varied patterns of partial and continuous reinforcement have received steadily increasing attention from psychologists. Although new and ingenious procedures are being devised for the study of this problem at the human level, the major emphasis continues to be on animal experimentation. A survey of partial reinforcement literature by Jenkins and Stanley (6) reveals that of the 30 studies reported during the years 1939-1949 few used humans as subjects. No significant change in this tendency is apparent from a review of the published research of the past five years. This fact has concerned some psychologists who have argued seriously for the greater utilization of human subjects in the testing of various aspects of learning theory. It is their contention that although constructs and laws coming from the animal laboratory have been shown to predict the behavior of infrahuman organisms with accuracy and usefulness, the time has come for a "validation" of some of these principles at the human level.

Although studies with humans in the area of partial reinforcement have supported, for the most part, results of animal research, a survey of the literature has disclosed two general characteristics of such experiments. First,

[1] This article is based on a dissertation submitted to the Iowa Child Welfare Research Station, State University of Iowa, in partial fulfillment of the requirements for the Ph.D. degree, June, 1953. The writer is indebted to Dr. B. R. McCandless under whose direction the research was performed, and to Drs. C. C. Spiker, A. Castaneda, I. E. Farber, and H. Bechtoldt for their helpful criticisms and suggestions.

[2] Now at the Wayne County Training School, Northville, Michigan.

very few have used the normal child as a subject. Apart from the fact that children can be conceived loosely as representing an "intermediary level" between inarticulate organisms and the ubiquitous college undergraduate, and, as such, constitute another valuable source of data, their use in learning studies seems vital for the development of a systematic approach in the field of child psychology, an area characterized until somewhat recently by a notable lack of sound, theoretically oriented research (7). Second, virtually no studies have been reported in which the effects of motivation have been intentionally introduced and systematically varied. Such situations as have been used [e.g., conditioning of the eyelid response (5), lever-pulling (3), and making predictions about the occurrence or nonoccurrence of a light (4)] would appear to lack "importance" to the subject, in the sense that nothing has been done to make them "ego-involving."

The present study was designed to investigate the effects on children's behavior of various reinforcement patterns in combination with motivational factors. Specifically, it concerns the effects of three schedules of reinforcement (100%, 50%, and 0%) at two levels of motivation (high and low) on the acquisition and extinction of a "certainty of prediction" response (CPR), that is, the degree of certainty with which a child makes a prediction concerning the occurrence of a success experience for himself.

The investigation can be conceived as having five stages. Stage one consists of determining an initial response measure (trial 1). During stage two (trials 1 to 10) changes in the CPR under high and low motivation conditions during 100%, 50%, and 0% reinforcement are determined. Stage three involves observation of the CPR at the end of trial 10. During stage four (trials 11 to 20) the effects of a 0% reinforcement series on the CPR are measured. And stage five (trial 20) consists of obtaining the terminal CPR.

The following hypotheses were formulated for testing:

1. Increase in mean CPR's during trials 1 to 10 would be significantly greater for the 100% group than either the 50% or 0% groups, with the 50% group evincing reliably greater increments than the 0% group.

2. Decrease in mean CPR's during trials 11 to 20 would be significantly greater for the 0% group than either the 100% or 50% groups, with the 100% group demonstrating reliably greater decrements than the 50% group.

3. Increase in mean CPR's during trials 1 to 10 would be greater for high in comparison to low motivation groups, except for the 0% group where no significant differences would manifest themselves.

4. Decrease in mean CPR's from trials 11 to 20 would be greater for low in comparison to high motivation groups, except for the 0% group where no significant differences would manifest themselves.

It should be noted with regard to the motivational variables that there is no systematic framework from which it is possible to derive hypotheses

concerning the performance of humans. Although some work has been done with adults (8), there is no general agreement at to whether increased motivational factors "help" or "hinder" a specific response. Hypotheses 3 and 4, then, were conceived purely as matters for empirical investigation.

Method

Materials

A series of "stimulus cards"—20 simple geometric designs drawn with a rule in black india ink on 4 by 5 in. index cards, a series of "comparison cards"—the same 20 designs drawn freehand in pencil on 4 by 5 in. yellow index cards, and a box of 15 red poker chips constituted the materials utilized in this experiment. The figures on the yellow cards were actually drawn by the experimenter, but were presented to each subject as the work of a "fourth (or fifth) grader just like yourself." The reason for this procedure is discussed below.

Subjects

The subjects used in this investigation were 120 white fourth and fifth grade children, 60 boys and 60 girls, with a mean CA of 10 years, 0 months, a mean IQ of 110.4, and a mean grade equivalent of 4.5. They were all obtained from one elementary school in Cedar Rapids, Iowa. Only "normal" children were included in the study; i.e., children who were, according to the principle's report, not undergoing any kind of psychological treatment procedure, were not crippled, etc. A superficial analysis of fathers' occupations revealed that the children came from predominantly middle-class homes.

Experimental Design

The subjects were assigned randomly to three major reinforcement groupings, and again randomized into motivational subgroups within each of the main groups. Group I received continuous positive reinforcement for trials 1 to 10, followed by a series of 10 nonreinforced trials; Group II received positive reinforcement only on trials 1, 4, 5, 7, and 10, followed by a series of 10 nonreinforced trials; and Group III received a continuous series of 20 nonreinforced trials.

Procedure

Each subject was seen individually in a private office and was read the following instructions:

> I'd like to play a game with you today. Other boys and girls have played this game with me and have liked it a lot. I'd like to see how you do at it.
>
> On this sheet of paper I'd like you to draw a design which I shall show you in just a moment. You can draw it anywhere at all on the page and you

can make it any size you want. Many other children in the fourth (fifth) grade have drawn these designs for me, and each time I present a new design to you on a white card like this (demonstrated with blank card), I'll show you on a yellow card like this (demonstrated) what kind of a job a fourth (fifth) grader just like yourself can do. Now, if you do a better job than this (pointing to yellow comparison card) I'll put a big plus mark on your paper like this (demonstrated), and if you do a poorer job I'll put a minus like this (demonstrated). And if you do just about as well I won't put anything on your paper. Get it? A plus mark for a better job, a minus mark for a poorer job, and nothing at all if you do just about as well. Draw these designs as quickly as you can, but work carefully since you won't be allowed to erase.

Before you begin, though, I'd like you to tell me what you honestly think you'll get for this first drawing, and to make it more fun for you I want you to use these chips to tell me. (Box and chips moved to center of table in front of subject.) Here is a pile of 15 chips which you are to use to tell me how certain you are of getting a *plus* mark on your paper. You can bet with as many or as few chips as you like to tell me how certain you are of getting a *plus*. For example, you can bet up to 15 of the chips if you're absolutely 100 per cent certain you're going to get a plus; or you can bet just *one* of the chips if you're not certain at all of getting it. Or you might want to bet somewhere in between, like 5, 6, 7, 8, or 9. Remember, the more certain you are of getting a plus on your paper the more chips you will want to bet. The less certain you are of a plus the fewer chips you'll want to bet. You can use from none at all to all 15 of them. You will make a bet every time before you draw. Count out your chips carefully, and put them in this box. Any questions? [3]

If the subject was part of a low motivation group he was then asked: "How certain are you of getting a plus mark on this first design?" If he was a high motivation subject he was read the following set of motivating (i.e., "ego-involving") instructions:

Now I'm interested in knowing who shall do better in your class, the boys or the girls. I know you'd like to see the boys (girls) come out on top, wouldn't you? All right, then, I want you to try to do the *very best* job you can. Also, I'm interested in knowing which fourth (or fifth) grade is best in Cedar Rapids. I'm sure you'd like your class to come out on top, so remember—do your very best! Now, how certain are you of getting a plus mark on this first design?

The subject then indicated the certainty of his prediction of receiving a plus mark for the first performance by counting out into the box whatever amount of poker chips he desired, after which a sheet of paper and the first pair of stimulus and comparison cards were presented. The comparison card was removed from the subject's vision after about a 4-second exposure,

[3] This betting method is a modification of one used by Castaneda (1). With this technique direct verbalization of prediction is avoided. Providing the child with chips that must be manipulated requires him to do some "subjective scaling" and careful thinking regarding each bet he makes. Nothing is won or lost with these chips. They were introduced solely as a technique of slowing down the subject so as to make him think more carefully about his response.

and then placed where only the experimenter could see it. The purpose of introducing such a card in the first place was to provide some "face validity" to the drawing game in general, and to prevent the subject from feeling the experimenter's evaluations were merely arbitrary.

After the subject had copied the stimulus figure the experimenter would pick up the paper, glance at it carefully, look over with a definite head movement to where the comparison card lay, glance back again at the subject's drawing and place either a large red plus mark on the paper or do nothing at all. A fresh sheet would then be put in front of the subject and the entire procedure begun again. While the child was busy copying the stimulus card, the experimenter would remove the chips counted out into the box and replace them in a single symmetrical pile with the rest of the chips on the table. After completion of the 20 trials each child was told he had done, on the whole, a very excellent job.

Results

The criterion measures used for testing the hypotheses of this study were the numbers of poker chips bet by a subject to indicate the degree of certainty with which he made a prediction of success for himself on each of the 20 trials constituting the experimental session.

Figure 1 represents schematically the results obtained on the three experimental groups. Each point on the curve is based on the mean number of chips bet (mean CPR's) by all 40 subjects in one group. An inspection of the curves at trial 1 reveal that mean CPR's do not differ significantly at that point, but that differences do manifest themselves with the advent of subsequent trials.

Trials 1 to 10

Using the technique of trend analysis for the data obtained for trials 1 to 10 the following results were obtained: (1) *Fs* of 25.15 for the 100% group, 3.93 for the 50% group, and 5.01 for the 0% group (all significant at less than the .001 level of confidence) revealed there were highly significant increments in mean CPR's from trials 1 to 10 under the three reinforcement schedules. (2) Tests for differences between pairs of curves revealed that significantly greater increments in mean CPR's were demonstrated under the 100% in comparison to the 50% or 0% schedules (significant at less than the .001 level of confidence), with no significant difference appearing between the 50% and 0% schedules. (3) No statistically significant differences appeared between the high and low motivation groupings.

Trials 11 to 20

It will be remembered that 10 successive nonreinforcements were presented to all subjects after they had been administered either 100%, 50%,

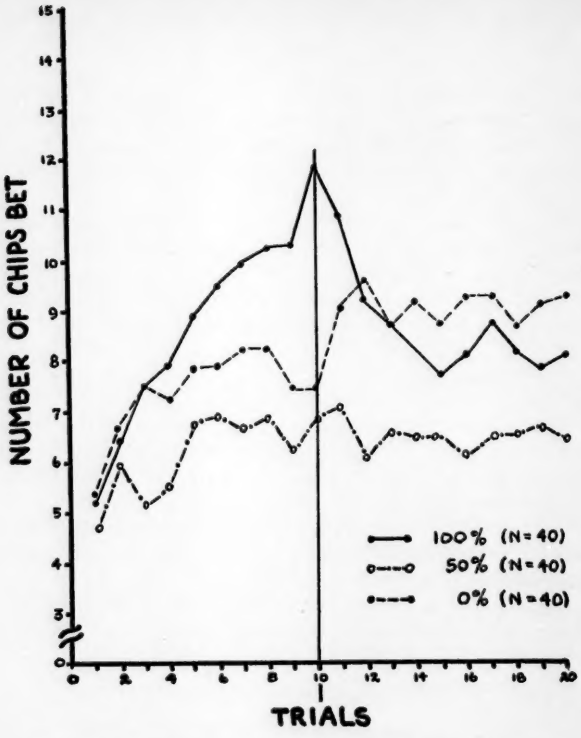

FIGURE 1—Mean certainty of prediction responses for all groups.

or 0% reinforcement for trials 1 to 10. Analysis of variance of the data revealed that the trials by reinforcement interaction (a test of whether the three groups manifested differential decrements over trials 11 to 20) barely missed statistical significance at the .05 level of confidence. This may have been due to the fact that the 50% and 0% groups, though evidencing no extinction trends in and of themselves, were contributing additional sources of variance to the total extinction data. Because of this insignificant T × R interaction it was not deemed feasible to test for differences between pairs of curves as had been done with the data obtained for trials 1 to 10. However, since an inspection of the empirical curve for the 100% group revealed a decided decline in mean CPR's from trials 11 to 20 whereas the 50% curve did not, it was decided to test for the presence of a 100% ex-

tinction curve. A sigma of the difference between the means of trials 11 and 20 netted a critical ratio of 3.48, which is significant at the .002 level of confidence. Thus, it appears a significant decrement does manifest itself for the 100% group, but this finding must be viewed with caution in light of the insignificant T × R interaction.

As with data for trials 1 to 10, comparisons of motivation results for trials 11 to 20 proved statistically insignificant.

Discussion

Results, then, supported that part of hypothesis 1 which posited that a group receiving 100% reinforcement would evidence greater increments in mean CPR's during trials 1 to 10 than would a group receiving either 50% or 0% reinforcement. This result conforms to those obtained from more conventional partial reinforcement studies using adult humans as subjects (2, 4, 5), which have demonstrated that acquisition of a response is directly proportional to the frequency of reinforced trials, proceeding, generally speaking, somewhat more rapidly under continuous than partial reinforcement.

On the basis of the findings of such studies it was further hypothesized that the group receiving 0% reinforcement during the first 10 trials would evidence statistically weaker increments in mean CPR's than would be the case with either the 50% or 100% groups. That this is not true is evident from an inspection of the curves in Figure 1. Not only was a significant increment trend demonstrated, but also there was no reliable difference between it and that of the 50% group.

Post hoc considerations of this somewhat paradoxical result lead to two tentative hypotheses, the first being that a state of anxiety was produced in these 0% subjects by the continuing series of nonreinforced trials (which may have been construed by the subjects as failure). If (*a*) anxiety is thought of as a drive, (*b*) it is considered that increased drive interacts with existing tendencies-to-respond so as to increase probability of strong response, and (*c*) the CPR can be thought of as a response of this nature, then the results for the 0% group can perhaps be considered from the point of view of such current psychological formulations as the Spence-Taylor type (9). Second, it might be that the increased CPR was due to some anxiety-reducing mechanism, perhaps of the "denial" variety. That is, it is possible these subjects tended to reject the reality aspects of the situation in which they were involved, steadily increasing their CPR's in spite of continued nonreinforcement. Because of a lack of opportunity to contrast positive and nonreinforcement, as was possible for the 50% group, the entire experimental situation was highly ambiguous for 0% subjects. Hence, for them an anxiety-reducing mechanism such as denial would be more "available" than for the 50% group subjects, who have experienced both positive and

nonreinforcement, and who would not be as likely to respond negatively to the reality aspects of the situation.

Results offered no statistical support to hypothesis 2 which predicted significantly greater decrements in mean CPR's during trials 11 to 20 for the 100% group than for either the 50% or 0% groups. There is a suggestion, however, from an inspection of the 100% and 50% curves and a consideration of the significant extinction test for the 100% data that results of these two groups are not completely deviant from the general conclusion reached in more conventional partial reinforcement research on human and animal subjects, namely, that a response will be more resistant to extinction if acquired under a schedule of partial rather than continuous reinforcement.

As was stated above, there was no systematic framework from which it was possible to predict the performance of high vs. low motivation groupings. The two motivation hypotheses in this study, therefore, were almost purely speculative, and were devised solely for purposes of empirical investigation. It will be remembered that an attempt was made to induce a high state of motivation in certain subjects by presenting a set of what were considered "ego-involving" instructions. It seems likely these instructions did not serve the purpose for which they were intended. Instructing a child at the beginning of the experimental situation to do his very best so that his sex will win top position in his class and his class top position among other similar classes in the city's schools apparently does not induce or sustain the desired state of motivation or drive. Were this experiment replicated it would perhaps be better first to match high and low anxiety groups on the basis of a technique such as the Iowa Anxiety Scale for Children and then induce motivation with a set of instructions geared to be more highly ego-involving. Such instructions could perhaps tell the child that the results of his performance during the experiment would be reported to his teacher, who would then incorporate them in a final grade for him.

To summarize, what has been demonstrated in this experiment is that Child A, instructed that three physical events are possible of occurrence during the course of the experimental situation (receipt of a plus, a minus, or nothing at all for a performance), will, significantly, become increasingly more confident that future success will occur as he receives a series of continuing successful experiences than will Child B, who is administered a number of success experiences interspersed with an equal number of nonsuccess experiences (defined as the receipt of neither a plus nor a minus), or Child C, who receives a series of continuous nonreinforcements. This is interpreted as a function of the greater frequency of successes experienced by Child A. Though not statistically significant, there is a tendency for certainty to diminish somewhat more rapidly for Child A than for B or C as each is administered the same number of continuous nonreinforcements after the first series of trials, with the latter two children manifesting no

significant decrements in response whatsoever. Child C, who does not experience success at all during the entire course of the experiment will, nonetheless, evidence steady increments in the certainty with which he predicts the occurrence of potential success for himself from trial 1 to 10. This is interpreted as a function of the presence of anxiety factors which are operating in this child.

Summary

This research was devised to investigate the acquisition and extinction of a response, namely, the degree of certainty with which a child makes a prediction concerning the occurrence of a success experience for himself, under three schedules of reinforcement (100%, 50%, and 0%) and two levels of motivation (high and low).

One hundred twenty normal fourth and fifth grade children were assigned randomly to three major reinforcement groups, and again divided into six motivational subgroups. Group I (100%) was administered a series of continuous positive reinforcements during trials 1 to 10, followed by a series of 10 nonreinforced trials; Group II (50%) received positive reinforcements on only five randomly selected trials during acquisition, followed by 10 nonreinforced trials; and Group III (0%) received a continuous series of 20 nonreinforced trials.

Each subject, seen individually, was presented with a "game" involving the copying of a series of 20 geometric designs. Before starting, he was informed that one of three possible events would occur after the completion of each individual drawing, the receipt of a plus, a minus, or nothing at all. Before copying each design he was instructed to "make a bet" from a pile of 15 poker chips to indicate how certain he was of the occurrence of a success experience for himself (i.e., the receipt of a plus). Any number of chips from 0 to 15 could be bet, zero indicating a success prediction of least and 15 a success prediction of most certainty. Nothing was won or lost with these chips; they were introduced solely as a nonverbal technique of obtaining a child's certainty of prediction responses (CPR's), and of getting him to do some "subjective scaling" regarding each bet. All subjects in "high motivation" subgroups were provided, additionally, with a set of specially prepared "ego-involving" instructions to induce a state of drive.

The criterion measures used in statistical analyses were the numbers of poker chips bet by each subject to indicate the degree of certainty with which he made a prediction of success for himself on each of the 20 trials constituting the experiment. In general, results yielded by the acquisition data are considered consistent with the findings of continuous and partial reinforcement research (using animals and adult humans as subjects) of somewhat more rapid acquisition of behavior under continuous than partial reinforcement. Though missing statistical significance slightly, results of the

100% and 50% extinction data are in the same direction as the general conclusion reached in these same studies, indicating greater resistance to extinction when behavior has been acquired under a schedule of partial as compared to continuous reinforcement. No significant differences between high and low motivation groups were demonstrated.

References

1. Castaneda, A. A systematic investigation of the concept expectancy as conceived within Rotter's Social Learning Theory of Personality. Unpublished doctor's dissertation, Ohio State Univer., 1953.
2. Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, 42, 1-5.
3. Grosslight, J. H., & Child, I. L. Persistence as a function of previous experience of failure followed by success. *Amer. J. Psychol.*, 1947, 60, 378-387.
4. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, 25, 294-301.
5. Humphreys, L. G. The effects of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid responses. *J. exp. Psychol.*, 1939, 25, 141-158.
6. Jenkins, W. O., & Stanley, J. C. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, 47, 193-234.
7. McCandless, B. R., & Rosenblum, S. Psychological theory as a determiner of experimental pattern in child study. *Rev. educ. Res.*, 1952, 25, 496-525.
8. Mosteller, F., & Nogee, P. Measurement of utility in a gambling situation. Reported in Jenkins and Stanley (6).
9. Spence, K. W., & Taylor, Janet A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, 42, 183-188.

AN OBJECTIVE METHOD FOR THE MEASUREMENT AND ANALYSIS OF CHILD-ADULT INTERACTION[1]

CLARK E. MOUSTAKAS, IRVING E. SIGEL, and HENRY D. SCHALOCK[2]
The Merrill-Palmer School

Interaction behavior is an important indicator of the underlying relationship between individuals. On the basis of overt behavior we often make inferences and interpretations and arrive at an understanding of its meaning and significance for the individuals involved. This kind of knowledge is possible when we have accurate presentations and descriptions of observable behavior.

How are we to collect the material from which more basic and dynamic understanding may be developed? What are the facts of interactional behavior which we can record, classify, and examine, and from which we may establish testable hypotheses? Observation appears to be the most appropriate technique for the descriptive phase of such a study. A method is needed by which reliable descriptive material on adult-child interaction can be obtained.

Our purpose here is to present a procedure for objective description and recording of adult-child interaction. Data are presented to show the usefulness of the procedure under three types of interpersonal situations. This

[1]The authors express their appreciation to The Merrill-Palmer School and its Director, Dr. Pauline Park Wilson Knapp, for making the research possible; to the parents who cooperated in the project and gave generously of their time and support; to Dorothy Haupt, Leader, Preschool Groups Service, who facilitated our approach to the children and assisted with the scheduling; to Drs. Chester D. Clapp and Esther B. Frankel, who introduced the parents to the observation procedures and laboratory; to Joan Godshalk, graduate student, who served as an observer and helped to establish reliability of the schedule; and to Dorothy L. Tyler, Editor of Publications, who prepared the manuscript for publication.

[2] Drs. Moustakas and Sigel are psychologists on the staff of the Merrill-Palmer School; Mr. Schalock, now on the faculty, Oklahoma A. & M. College, was a graduate student, 1953-54.

procedure involves the use of a predetermined set of general purpose observational categories which can be applied to different types of adult-child interaction.

A major source of suggestions for these categories was the excellent study of Barbara Merrill Bishop (2). Her schedule for studying mother-child social interaction was used as a point of departure. We then made refinements, clarified category definitions, and added numerous categories to make our schedule more inclusive, detailed, and widely applicable. The additional categories were in part selected from other studies using a system of behavioral categories (3, 4, 5, 6, 7, 8) and in part were based on the authors' experience. A scale for recording anxiety and hostility was added. Further revisions, suggested by experience in using the schedule to observe mother-child and therapist-child interaction, are incorporated in the final schedule.

Selection of Categories

Selection is extremely important in establishing a set of categories for observation. It is often based on the nature of the problem and on a specific theoretical framework. Though the present schedule was constructed on an empirical basis, in order to develop a schedule that would have wide application, certain theoretical assumptions are nevertheless inherent in the schedule. The primary assumption is that interactional behavior is manifold and composed of various types of observable behaviors. Another assumption is that it is possible to construct overt behavior categories that will not only have wide applicability for various kinds of adult-child interaction but will also be of use and value to those holding different theoretical points of view.

The criteria used in selecting the categories in the schedule were: comprehensiveness, relevance and meaningfulness, and ease of identification.

Comprehensiveness. This criterion refers to the categories *in toto,* that is, the question of whether they embrace a relatively inclusive number of adult-child interactional behaviors. Review of the literature, together with the experience of observers and writers, was the basis for ascertaining what categories should be selected to cover the major aspects of such interactional behavior.

Relevance and meaningfulness. Certain arbitrary decisions were made as to the contribution certain categories made in furthering our understanding of adult-child interaction processes. Here our own theoretical biases concerning adult-child roles, expectations, and consequent behaviors influenced our decisions and are evident in the specific categories selected.

When an adult interacts with a child he may influence the child in many different ways to bring about the kind of behavior he wishes. He uses the type of behavior appropriate to his goal. An adult has a certain

status position derived from his maturity, size, and function. As a mature, socialized individual he can use his position to frustrate the child's needs or to gratify them. He can initiate certain kinds of behavior in the child and ignore others. Certain categories were derived on the basis of these influence characteristics of the adult.

Although a child has relatively limited size, maturity, and socialization, he can also influence the adult. He can frustrate the adult, intentionally or otherwise; he can also direct and gratify the adult. Thus, though not powerful, not mature, and not operating with the same general wherewithal, the child can influence adult behavior. These child influence expressions are also a source of categories in the schedule.

In order to keep the schedule of manageable length, we included a category only when it offered information that would help us more adequately to understand our own conception of adult-child interaction.

Ease of identification. This criterion can be applied only when categories can be precisely, clearly, and adequately explained. That is, trained observers must find the relevant behaviors relatively easy to identify and label. For this reason certain details of behavior were omitted from the schedule. The greater the ease with which observers could label the observed behavior, the greater the chance for observer agreement.

The Schedule: Discussion and Analysis

Schedule categories for observing and scoring adult behavior and child behavior are mainly the same.[3] Definition, illustration, and discussion as applied to child and adult are therefore presented together in this section under the various category headings. Table 1 presents a complete list of categories, with scoring symbols.

Attention Categories

Nonattention. *A* (adult) directs his attention to something other than *C* (child), or *C* directs his attention to something other than *A*.[4] Ex: *A* reads a book or leaves the room for a personal reason. Ex: *C* plays quietly with back to *A* or stares out the window. Nonattention may be accompanied by humming, singing, or talking, but this behavior is not apparently directed toward *C* or *A*. No interactive characteristics are observable.

[3] Only two categories differ: category *z*, Interpretation, broadly defined in the child categories and defined in greater detail and more specifically in the adult categories; *x*, Threat of Attack, is omitted from the adult categories.

[4] After reliability was established, a broad rating scale was constructed to provide a general measure of the frequency of negative and positive expressions in the noninteractive behavior of *A* and *C*. The rating is based on tonal quality, choice of gestures, words, and the content of the expression, as follows: 0 Behavioral expression appears neither positive nor negative. 1 Behavioral expression is positive, e.g., "What a lovely doll!" (*C* caresses the doll.) 2 Behavioral expression is negative, e.g., hitting or kicking objects, tenseness of facial muscles while drawing, etc.

Attentive Observation. *A* noticeably directs his attention to *C* and/or *C*'s activity by silently watching; or *C* directs his attention to *A* in this way. Ex: *A* watches *C* as *C* plays in the sand. Ex: *C* watches *A* as *A* draws a picture.

Recognition. *A* responds to *C*'s stimulation in a way that indicates definite awareness of *C*'s statement or activity; or *C* responds to *A*'s stimulation in this way. Ex: Mm hm. Ex: Yes, I see. Ex: (*A* listens and generally attends to *C* as *C* tells about an experience.)

These two categories, Attentive Observation and Recognition, convey interest and concern on the part of adult or child. Attentive Observation may have an important influence on the behavior of child or adult, but this influence is not as clear as that associated with Recognition. While *A* is attentive to *C*, *C* may be unaware of *A*; but when *A* recognizes *C*, this recognition must be based on a stimulus directed to *A*. As defined in this schedule, Recognition is always interactive, while Attentive Observation may be unidirectional.

Statement of Condition or Action. A comment describing an existent state, situation, or action. Ex: It's hot in here. Ex: I'm afraid. Ex: I'm going now. Though such statements of condition or action may have stimulus properties, from the observer's point of view they apparently do not influence behavior or elicit particular responses.[5]

Joint Participation in Activity. *A* and *C* are mutually engaged in an activity. Ex: *A* and *C* are reading, working with puzzles, singing, playing, drawing, or painting together.

Offering Information. *A* or *C* offers knowledge or guidance. (1) Information is offered verbally. Ex: *C*: Why won't this open? (Referring to cash register.) *A*: You have to press the keys hard to make it open. Ex: *C*: You have to unsnap my boots before you can take them off. (2) Information is offered by means of demonstration. Ex: (*C* is apparently trying to make wheels with tinker toys and is having difficulty.) *A*: Let me show you how to do one. (*A* proceeds to demonstrate how to make a wheel.) Ex: *C*: Let me show you how to work this puzzle. (*C* demonstrates.) (3) Information is offered through the combined methods of demonstration and explanation. Ex: (*C* is having difficulty working puppets. *A* takes puppet, shows *C* how to work it.) *A*: You put your thumb in here and work it like this. Ex: (*A* asks *C* how the rifle works.) *C*: Here's how you shoot it. You keep pumping it until the ball flies out.

Giving Help.[6] Physical aid is given by *A* or *C*. Ex: (*C* is trying to hook the cars of a train together but is having no success. *A* holds cars while *C* hooks them together.) Ex: (*C* holds finger on string while *A* ties a knot.)

---

[5] After reliability was established, and on the basis of further experience, the category Statement of Condition or Action was subdivided as follows: b_1 or b_4 Statement involving a situation external to the individual. Ex: I went to the circus yesterday. Ex: It's cold outside. b_2 or b_5 Statement involving an emotional state of the person. Ex: I'm afraid. Ex: I hate my father. b_3 or b_6 Statement of action or proposed action. Ex: I'm making a scrapbook. Ex: I'm leaving now.

[6] After reliability was established, and on the basis of further experience, the category Giving Help was subdivided as follows: h_1 or h_4 *A* helps *C* in completing a task; or *C* responds to *A* in this way. Ex: *C* helps *A* put tinker toys away. Ex: *A* holds finger on string while *C* ties a knot. h_2 or h_5 *A* helps *C* by completing a task for *C*; or *C* responds to *A* in this way. Ex: *A* makes tinker toy objects for *C*.

These two categories, Offering Information and Giving Help, both involve assistance to the child or adult. However, a distinction may be made. Information involves a kind of teaching, through demonstration, explanation, or the stating of a principle or fact. Helping, as used here, means actually doing or helping to do the task at hand.

Reassurance. Comfort or encouragement is offered by *A* or *C*. Ex: *A*: Don't feel too bad. Most children find it difficult to work that puzzle the first time. Ex: *A*: It doesn't matter that you spilled some paint on the floor. It can be easily washed. Ex: (*A* drops a piece of doll furniture, which breaks.) *C*: Don't worry, I've got some other furniture to play with.

This category, Reassurance, differs from the categories Giving Help and Offering Information in not involving information or assistance relevant to solving a problem. It is a method of offering assistance through emotional support.

Stimulus Categories

Several categories involve an attempt to elicit a particular kind of response or behavior by *A* or *C*. These stimulus categories are discussed next.

Seeking Information. *A* or *C* asks for information. (1) Personal information is sought, i.e., information about parents, sibling or peer relationships, feelings toward school, etc. Ex: *A*: How is everything in school these days? Ex: *A*: Does your little brother break your toys? Ex: *C*: Mommy, are you mad at Daddy? (2) Information is sought concerning the ongoing activity. Ex: *A*: Do you like that puppet more than the others? Ex: *C*: What is this thing?

Seeking Help.[7] *A* or *C* asks for help. (1) *A* seeks assistance in completing a task; or *C* responds to *A* in this way. Ex: Help me carry these chairs. (2) *A* requests *C* to complete a task for him; or *C* responds to *A* in this way. Ex: *C*: Will you pound this in for me? I can't do it right.

Seeking Reassurance. Comfort or encouragement is sought by *A* or *C*. Ex: Are you sure you're not angry with me? Ex: Am I doing this the way I should?

Seeking Recognition. Attention to activities, productions, or statements is sought by *A* or *C*. Ex: See what I've made? Ex: Look what I'm doing.

Seeking Praise. Desire for praise is indicated by *A* or *C*. Ex: Don't you think this is a nice picture? Ex: How do you like the way I fixed your hair?

Seeking Affection. Expressions of love or tenderness are sought by *A* or *C*. Ex: Come and kiss me. Ex: Hold me on your lap.

Seeking Reward. Special privileges or benefits are sought by *A* or *C*. Ex: I've been a good boy all day. Will you give me some candy? Ex: If you let me go downtown, I'll bring you back an ice cream cone.

Seeking Permission. Consent is sought by *A* or *C* for a proposed activity. Ex: Is it all right if I leave for a minute? Ex: Could I play with all these dolls?

Orienting and Directing Categories

Another group of related categories in the schedule are Orienting, Directing, Restricting, and Forbidding. Orienting sets the stage and usually permits a wide

[7] Revised form given. The earlier form read: *A* or *C* requests physical assistance. This may be done either verbally or nonverbally. Ex: Help me carry these chairs. Ex: I can't pound this right. Will you hold this peg for me?

range of behavior. Directing tends to focus attention or activity on specific behavior or in a particular direction, and is narrowly oriented. Restricting and Forbidding are also focused on particular behavior in a specific way, narrowly limiting or completely stopping the behavior.

Orienting. Limits, boundaries, and roles are indicated by *A* or *C*.[8] (1) Boundaries of the situation are indicated by *A* or *C*. Ex: You may do whatever you like here. (2) *A* or *C* structures time. Ex: You have only about five more minutes left to play. Ex: You can stay only a few more minutes. (3) Roles are structured by *A* or *C*. (*a*) Role is indicated on the basis of status or power by *A* or *C*. Ex: C: How do you know I'll get sick if I eat candy? *A*: Older people always know these things. Ex: *C*: Mommy, you be the sheriff and I'll be the cowboy. Ex: *A*: You'll have to ask daddy. He makes the rules. Ex: *C*: I'm the smallest, so I get my turn first. (4) Roles are indicated by leaving responsibilities to *C* or *A*. Ex: I can't tell you what to do, you must decide for yourself. Ex: How do you think it should be done? Ex: You can use these things any way you want.

Directing. *A* attempts to influence *C* by command or suggestion; or *C* attempts to influence *A* in this way. (1) *A* or *C* uses verbal or nonverbal suggestion. Ex: You could play with the darts if you wanted to. Ex: Would you like to sit over here? Ex: (*C* is hitting table top with hammer. *A* places pounding block on table.) (2) *A* or *C* gives a command accompanied by a threat. Ex: Clean up this mess or you can't play in here any more. Ex: If you don't bring me that knife, I'll throw this at you. (3) *A* or *C* gives a command. Ex: Put the gun on the table. Ex: Get me that book.

Restricting. *A* attempts to modify *C*'s behavior by reducing the intensity, speed, manner of executing, etc., but apparently does not intend to stop the activity completely; or *C* attempts to modify *A*'s behavior in this way. (1) *A* attempts to restrict *C*'s ongoing activity by warning; or *C* responds to *A* in this way. Ex: Be careful, or you will fall and hurt yourself. Ex: You hadn't better come any closer, or I might splash paint on you. (2) *A* attempts to restrict *C*'s ongoing activity, while offering an explanation; or *C* responds to *A* in this manner. Ex: Don't splash the water so high, Jimmy, or you'll get your clothes wet. Ex: Don't open the window so high, or I'll get cold. (3) *A* attempts to restrict *C*'s ongoing activity by threat of bodily punishment; or *C* threatens *A* in this way. Ex: It's all right to paint, but if you spill any more on the table you'll get spanked. Ex: Don't talk so loud, or I'll slap you. (4) *A* attempts to restrict *C*'s ongoing activity by threat of loss of affection; or *C* attempts to restrict *A* in this way. Ex: I can't be nice to you when you yell so loud. Ex: I won't sit on your lap if you don't read slower. (5) Restriction of the ongoing activity is attempted by *A* or *C* through a direct statement. Ex: I don't think you should pound the table quite so hard. Ex: Don't make me go so fast. (6) *A* recognizes or clarifies *C*'s feelings while attempting to restrict *C*'s ongoing activity; or *C* responds to *A* in this way. Ex: I know you'd like to eat all those cookies, but you may only have two. Ex: I know you're very mad at me, Mommy, but don't push me so hard. (7) Restricting the ongoing activity by limiting the use of objects is attempted by *A* or *C*. Ex: You may fill that truck only once more.

[8] After reliability was established, and on the basis of further experience, all subcategories of Orienting were included in the observation schedule for the child.

Ex: You may use only four blocks to build the house. (8) Restricting the ongoing activity by physical restraint is attempted by *A* or *C*. Ex: (*C* attacks by hitting. *A* attempts to regulate the force of the blow.) Ex: (*A* slaps *C*. *C* breaks the force of the blow.)

Forbidding. An ongoing activity is interfered with by *A* or *C* with the apparent intent of stopping it completely. (1) *A* attempts to stop *C*'s ongoing activity while offering an explanation; or *C* attempts to stop *A* in this way. Ex: I can't let you paint on the walls, because the paint won't come off. Ex: Stop putting my toys away. I want to play with them. (2) *A* attempts to stop *C*'s ongoing activity while offering an alternative; or *C* responds to *A* in this manner. Ex: I can't let you shoot me in the face, but you can shoot me in the body if you want to. Ex: You can't have that chair, but you can sit on the other one. (3) Feelings are recognized and clarified by *A* or *C* while an attempt is made to stop an ongoing activity. Ex: I know you feel like hurting me, but I can't let you do that again. Ex: I know you don't like me telling you what to do, but stop saying what I say. (4) *A* or *C* attempts to stop an ongoing activity by threat of bodily punishment. Ex: If you don't stop making so much noise, I'm going to spank you. Ex: You shut up, or I'm going to hit you. (5) *A* attempts to stop *C*'s ongoing activity by threat of loss of affection; or *C* responds to *A* in this manner. Ex: I won't like you any more if you don't stop splashing me with water. (6) *A* or *C* attempts to stop an ongoing activity by means of a directly phrased statement. Ex: I will not permit you to hit me with that again. Ex: Stop that! (7) *A* attempts to stop *C*'s ongoing activity by proceeding to remove objects from *C*; or *C* responds to *A* in this manner. Ex: *C* is pounding on window with toy hammer. *A* takes hammer from *C*. Ex: *A* is writing on a note pad; *C* takes pencil from *A*. (8) *A* or *C* attempts to stop an ongoing activity by employing physical barriers. Ex: (*A* closes door when *C* tries to leave room.) Ex: (*C* blocks *A*'s path when *A* attempts to get coat and hat.) (9) *A* or *C* uses physical restraint in attempting to stop an ongoing activity. Ex: (*C* slaps at *A*. *A* hold's *C*'s arms, stopping the attack.) Ex: (*A* starts to leave room. *C* stops *A* by pulling on coat.) (10) *A* attempts to stop *C*'s ongoing activity by warning; or *C* attempts to stop *A* in this way. Ex: If you don't stop swinging those scissors, you'll hurt yourself. Ex: If you don't stop reading, you'll get a bad headache.

Criticism, Discipline, Rejection Categories

The next five categories in the schedule are Criticism, Disciplinary Action, Physical Attack, Threat of Attack, and Rejection. Threat of Attack, when it occurs without apparent provocation, is used only in observation of the child's behavior. Threats of attack with apparent provocation are scored in the specific categories involved; i.e., restricting or forbidding.

Criticism. *A* gives a negative evaluation of *C* or *C*'s productions; or *C* responds to *A* in this way. (1) *A* makes a negative evaluation of *C* by employing shame, ridicule, or disparaging remarks; or *C* responds to *A* in this manner. Ex: Don't be so stupid. Ex: Why don't you act your age? Ex: Damn you! Ex: Ha! Ha! Spilled milk all over yourself. (2) *A* makes a negative evaluation of *C*'s productions by employing shame, ridicule, or disparaging remarks; or *C* makes negative evaluation of *A*'s productions in this way. Ex: That's a sloppy way to paint. Ex: Aw, this isn't any good! You

don't know how to make pictures like my teacher does. (3) *A* makes a negative evaluation of *C* and *C*'s productions by employing shame, ridicule, or disparaging remarks; or *C* responds to *A* in this way. Ex: When you talk like that, you are acting like a baby.

As used in the schedule, Criticism always involves some kind of negative evaluation; therefore, constructive criticism is not included in this category. When constructive criticism occurs in a child-adult interaction, it may involve giving information, suggesting, directing, etc., and is scored in the specific category involved.

Disciplinary Action. *A* attempts to inflict pain on *C* as a response to an indicated undesirable act of *C*; or *C* responds to *A* in this way. The concept of punishment here usually includes moral, ethical, or teaching values, but it may involve solely the intent to inflict pain or hurt. (1) *A* punishes *C* by isolation; or *C* punishes *A* in this way. Ex: I've told you not to do that before. Now you go to your room and stay there. (2) *A* punishes *C* by deprivation of physical objects, food, privileges, or restriction of space; or *C* responds to *A* in this way. Ex: Just for that, you can't play with the gun any more. Ex: I told you not to read. (*C* takes paper from *A* and hides it.) (3) *A* punishes *C* by destroying or attempting to break objects in response to *C*'s apparently irritating behavior; or *C* responds to *A* in this way. Ex: (*C* breaks toy after seeing *A* kiss baby.) (4) *A* punishes *C* by means of corporal punishment; or *C* punishes *A* in this manner. Ex: (*C* sasses *A*.) *A*: That does it. (*A* slaps *C*.) Ex: Just for that I'm going to spank you. (*C* proceeds to spank *A*.) (5) *A* punishes *C* by means of deprivation of love or affection; or *C* punishes *A* in this way. Ex: Since you won't sit still, I'm not going to hold you. Ex: Since you won't let me go, I'm not kissing you good-by.

Physical Attack. *A* strikes *C*, or *C* hits *A* without apparent provocation. Ex: *C* kicks *A*.

Threat of Attack. *C* threatens to attack *A* without apparent provocation. (1) *C* verbally threatens to attack *A*. Ex: I'm going to hit you. (2) *C* threatens to attack *A* by gesture. Ex: (*C* approaches *A*, holding a toy hammer in striking position.) (3) *C* both verbally and gesturally threatens *A*. Ex: I'm going to shoot you with this rifle. (Aims air rifle at *A* and moves toward *A*.)

Rejection. *A* indicates nonacceptance of *A* and/or *C*'s stimulation, either verbally or gesturally; or *C* responds to *A* in this way. (1) *A* changes the subject or interrupts *C* with irrelevant conversation; or *C* responds to *A* in this manner. Ex: (*C* asks *A* to help him build a swing with tinker toys. *A* changes the subject, saying, "You haven't painted today yet.") Ex: *A*: Pick up the blocks. Ex: *C*: It's cold in here, isn't it? (2) *A* denies the validity of *C*'s statement or action; or *C* responds to *A* in this manner. Ex: *C*: I'm a good boy. *A*: No, you are a bad boy. Ex: *A*: That's a house. *C*: No, it's not a house. Ex: (*A* shakes head negatively.) (3) *A* rejects *C* as a person; or *C* rejects *A* as a person. Ex: I don't want anything to do with you. Get away from me. Ex: I hate you. Ex: (*C* asks *A* to play with him. *A* walks away.) (4) *A* makes a direct appeal to *C*, and *C* ignores or evades him; or the appeal is made by *C* and *A* ignores or evades him.[9]

[9] This category was originally classified as Nonrecognition, and later included in the group of rejection categories. Nonrecognition is coded as n in Table 2.

Ex: (*C* looks at *A* and asks, "Do you like me?" *A* does not respond.) Ex: *A*: Don't you think this is a nice picture I have made for you? (*C* does not respond.)

Approval or Reward Categories

Four approval or reward categories are included in the schedule. They are Permission, Praise, Affection, and Reward.

Permission. *A* consents to *C*'s proposed activity; or *C* consents to *A*'s proposed activity. (1) Permission is granted with conditions. Ex: Yes, you may use the paints if you put on an apron. Ex: Yes, you may go if you come back in a few minutes. (2) Permission is granted without condition. Ex: Yes, you may have another sheet of paper. Ex: Yes, you may leave now.

Praise. *A* expresses approval of *C* and/or *C*'s productions; or *C* responds to *A* in this way. Ex: That's a fine picture you've made. Ex: You can sure write well.

Affection. *A* gives overt expressions of love or tenderness to *C*; or *C* responds to *A* in this manner. (1) Affection is expressed by physical means. Ex: (*A* rubs or pats *C*'s head.) Ex: (*C* puts arms around *A*.) (2) Affection is expressed verbally. Ex: I love you very much. Ex: I like you. (3) Affection is expressed by physical and verbal means. Ex: *A* hugs *C* and says, "Mm, I like you very much." Ex: *C* kisses *A* and says, "I love you."

Reward. *A* extends special privileges or benefits to *C* as a response to *C*'s behavior; or *C* extends special privileges or benefits to *A*. Ex: For being so good today, you may stay ten minutes longer. Ex: For helping me with this picture, I'm going to give it to you.

Cooperation, Compliance, Noncooperation

These are essentially response categories. Cooperation, which has no coercion aspects, involves deliberate participation, with enthusiasm and willingness. Compliance involves submission and implies coercion and unwillingness. It may or may not be destructive or unreasonable. Noncooperation involves a refusal to participate, an action that goes against the stated request or apparent expectation of child or adult.

Cooperation. *A* responds to *C*'s comments, suggestions, or requests with apparent interest and willingness; or *C* responds to *A* in this manner. (1) Casual or straightforward participation. Ex: O.K., I'll play darts with you. Ex: O.K., I'll put the toys away. (2) Enthusiastic participation. Ex: I would love to play house with you. What could I do? Ex: O.K., that's a very good idea. I'll do it now.

Compliance. Responds to commands, suggestions, or requests with apparent reluctance or indifference. Ex: Oh, all right, I'll do it, but I don't want to. Ex: All right, but I'd rather play with the gun.

Noncooperation. Refusal to accept commands, suggestions, or requests. (1) Refusal to participate accompanied by an explanation. Ex: I won't leave yet. I haven't finished building the house. Ex: I won't play in the sand with you, because I don't want to get my clothes dirty. (2) Refusal to participate accompanied by an alternative. Ex: I won't play in the water with you, but I'll hand you the boats. Ex: I don't want to put the gun on that table, but I'll put it on the ledge. (3) Simple refusal to participate. Ex: No! Ex: I'm sorry, but I won't

do that. (4) Strongly expressed refusal to participate. Ex: No! I certainly will not. Ex: No! No! No!

Ambivalent statement. *A* expresses conflicting ideas, philosophy, or attitudes; or *C* does so. Ex: I don't know whether to let you do that or not. Ex: I like it here, but I'm afraid.

Interpretation

Interpretation, the final category in the schedule, is broadly defined for use in observations of child behavior. Eight separate subcategories are included in the adult schedule. In general, an interpretive response is one which in some respects moves away from expressed statements. As the adult response uses less and less of what the child immediately expresses, the response is more and more removed from the child's awareness and is therefore considered more interpretive. Probably correspondence of phenomenological fields is never identical, and therefore some aspect of interpretation is involved. Individuals respond to communication in terms of their own cognitive organizations, selecting aspects of the stimuli complex as the basis for response.

Interpretation.[10] *A* comments on *C*'s behavior by means of evaluation or analysis, selecting and emphasizing certain aspects of it in terms of theoretical concepts, beliefs, or assumptions, or from knowledge of past history; or *C* responds to *A* in this way. (1) *A* selects an aspect of *C*'s remark and restates the content of the remark. Ex: *C*: That's a car. This is a truck. Here's the baby. *A*: That's the baby, huh? (2) *A* verbalizes the content of *C*'s motor behavior. Ex: (*C* makes four drawings.) *A*: I see that you have made four drawings. (3) *A* restates feelings of verbal statements. Ex: *C*: I don't like that paint. *A*: You just don't like it at all. (4) *A* recognizes feelings in motor behavior. Ex: (*C* pours paint on paper and splashes it around.) *A*: You feel like messing it all up. (5) *A* attempts to clarify *C*'s expressed feelings by elaborating *C*'s statements. Ex: (*C* pours paint from can into container, using more and more paint.) *C*: I love to mix this. *A*: Maybe you feel like putting the whole can in. (6) *A* attempts to clarify *C*'s feelings when the feelings are not immediately verbalized or indicated in *C*'s motor behavior. Ex: (*C* takes red paint and begins to paint mother doll's hair, clutching figure firmly, with intent expression on face.) *A*: Maybe you just feel like smearing her face. (7) *A* attempts to clarify *C*'s feelings, which are not immediately verbalized or indicated in *C*'s motor behavior, by pointing out a temporal relationship between an immediate expression of *C* and some event in the past. Ex: *A*: Yesterday you really knocked the big brother around, but today you seem afraid. Perhaps you would like to hurt him, but are afraid of what he might do to you. Ex: (*C* drinks from nursing bottle.) *A*: You are drinking from the nursing bottle the way you did when you were a baby. It makes you feel good now like you felt when you were a baby. (8) *A* responds to *C* by giving his account of reality to *C*, by explanation and/or translation of symbolic behavior. Ex: *C*: I'm just scared and don't know what it is. *A*: I know. You are afraid because you are alone in this room with a man you do not know and whom you cannot yet trust. Ex: (*C* puts two swings together and points out the inner part, which hangs down and swings.) *C*: Look how it dangles and bumps. (*A* points to dangling swings and says:) Those are two

[10] The subcategories apply only to observations of adult behavior.

people, your Daddy and Mommy, bumping their thingummies [*C*'s word for genitals] together.

Anxiety-Hostility Ratings

The ratings of anxiety and hostility attempt to measure the intensity of two types of emotional concomitants of behavior. Each rating is based on the style in which the behavior manifests itself. The observer judges the emotional expressions involved by utilizing various clues, such as facial expression, feeling tone of language, body tension, random body movements, tenseness, and tremors. With these manifestations as his bases, the observer judges intensity by ascribing a rating to the specific behavior.

The following ratings are used:

0 = Little or no anxiety or hostility apparent.
+ or − = Indication of some anxiety or hostility.
++ or −− = Indication of much anxiety or hostility.

Such a rating accompanies each of the interactive categories in the schedule.

Conditions and Procedures for Observation

Laboratory Observations

Laboratory observations were made of mother-child and therapist-child situations in a room set up primarily for play therapy situations, and equipped with a variety of toys, materials, and equipment (9). The observers were seated in an observation booth, wired for sound, behind a one-way vision mirror. The therapist knew that he was being observed, but the children were not aware of the observers. In the mother-child situation, both participants knew they were being observed. The mothers were told that the purpose of the visit was to help us work out methods for observing parents and children in interaction.

The therapist-child situation was based on principles and concepts of child-centered play therapy. The therapist explained to the child that he was free to use the materials in the playroom in any way he wished during the 45-minute session. Mother and child were instructed that they could do whatever they wished while in the playroom, and that they would be notified when the 45-minute observation was over.

Home Observations

Three home observations were made in order to achieve observer agreement and to allow sufficient time for mother and child to become accustomed to the observer's presence (1). Initial practice home observations were made in a family well known to both observers. After the procedures were somewhat familiar, contacts were made with other families. A standardized procedure was employed when a family was to be observed. In a preliminary interview in the home the purpose of the observation was

explained and rapport was established with mother and child. The observer also had an opportunity to see the home and so knew in advance where the child usually played, where toys were kept, and what space, if any, was forbidden to the child. An appointment was then made for the initial observation, and the mother was told that during the observation she should proceed with her usual tasks.

Certain home observation principles were established: The observer was to remain close enough to the mother-child interaction to see and hear what occurred, yet as much apart from it as possible. The observers tried to maintain a free, noninteracting relationship with mother and child, especially during the time of actual observation. If the child made friendly or attention-getting overtures, the observer was to recognize him but to pay no further attention. As a rule, the child needed only this brief recognition to enable him to begin or return to an activity. If the child persisted, the observation was stopped and the child was told the observers were working and would talk to him when they had finished. The mother was usually helpful at such a time. In no instance did a child continue such overtures throughout an observation period.

The positions taken by the observers depended upon the setting, but they strove to maintain unobtrusive positions. If mother and child were in a relatively small room, the observers usually stationed themselves in or near the doorway, where they were apart from the immediate scene, could move out of the way if mother or child left the room, and could shift their position if the scene of interaction changed. When mother and child were in a larger room, the observers could again operate from the doorway or assume an unobtrusive position on either side or end of the room. When the interaction occurred outside the room, the observers remained in the background, as far as possible.

Whenever mother and child were out of each other's sight, the observers remained with the mother. This procedure was found to increase opportunities for observing mother and child interaction, since the child would return his attention to the mother, but the child, when the observer stayed with him, often interrupted the observer and made fewer contacts with the mother. This finding suggests again the value of a series of home observations in reducing the influence of the observer on the kind and quality of mother-child behavior.

Recording Principles and Procedures

The behavior of child and adult is recorded individually for each 5-second period. Recordings are made in terms of category code letters (Table 1) and entered on a prepared scoring sheet, designed for 16 minutes of continuous recording, each square representing 5 seconds of time. A stop watch is used.

TABLE 1

A SCHEDULE FOR OBSERVING AND SCORING CHILD-ADULT INTERACTION

Category	Scoring Symbol: Adult	Scoring Symbol: Child*
Nonattention	o	*o*
Neutral behavioral expression	o_0	o_0
Positive behavioral expression	o_1	o_1
Negative behavioral expression	o_2	o_2
Attentive observation	c	*c*
Recognition	e	*e*
Statement of condition or action		
Statement of situation	b_1	b_1
Statement of emotional state	b_2	b_2
Statement of action	b_3	b_3
Joint participation in activity	p	*p*
Offering information		
Verbal information	t_1	t_1
Information by demonstration	t_2	t_2
Information by demonstration and explanation	t_3	t_3
Giving help		
By assisting in completing task	h_1	h_1
By completing task	h_2	h_2
Reassurance	a	*a*
Seeking information		
Personal information	t_1—	t_1—
Impersonal information	t_2—	t_2—
Seeking help		
By requesting assistance	h_1—	h_1—
By requesting completion of task	h_2—	h_2—
Seeking reassurance	a—	*a*—
Seeking recognition	e—	*e*—
Seeking praise	m—	*m*—
Seeking affection	l—	*l*—
Seeking reward	g—	*g*—
Seeking permission	Q—	*Q*—
Orienting		
Boundaries	s_1	s_1
Time	s_2	s_2
Roles		
By status, power, age	s_{3m}	s_{3m}
By indicating responsibility	s_{3n}	s_{3n}

(*continued on next page*)

TABLE 1 (*continued*)

Category	Scoring Symbol *Adult*	*Child**
Directing		
Suggestion	d_1	d_1
Command and threat	d_2	d_2
Command	d_3	d_3
Restricting		
By warning	r_1	r_1
By explanation	r_2	r_2
By threat of bodily punishment	r_3	r_3
By threat of loss of affection	r_4	r_4
By direct statement	r_5	r_5
By recognition and clarification	r_6	r_6
By limiting the use of objects	r_7	r_7
By physical restraint	r_8	r_8
Forbidding		
By explanation	f_1	f_1
By offering alternative	f_2	f_2
By recognition and clarification	f_3	f_3
By threat of bodily punishment	f_4	f_4
By threat of loss of affection	f_5	f_5
By direct statement	f_6	f_6
By removing object	f_7	f_7
By physical barriers	f_8	f_8
By physical restraint	f_9	f_9
By warning	f_{10}	f_{10}
Criticism		
Of person	n_1	n_1
Of person's production	n_2	n_2
Of person and person's productions	n_3	n_3
Disciplinary action		
By isolation	i_1	i_1
By deprivation of objects or space	i_2	i_2
By attacking person's objects	i_3	i_3
By corporal punishment	i_4	i_4
By deprivation of love or affection	i_5	i_5
Physical attack	v	*v*
Threat of attack		
By verbal expression		x_1
By gesture		x_2
By verbal and gestural means		x_3
Rejection		
By changing subject	w_1	w_1
By denying validity of statement	w_2	w_2
As a person	w_3	w_3
By ignoring or evading	w_4	w_4

(*continued on next page*)

TABLE 1 (*continued*)

Category	Scoring Symbol	
	Adult	*Child**
Permission		
With qualifications	Q_1	Q_1
Without qualifications	Q_2	Q_2
Praise	m	*m*
Affection		
By physical means	l_1	l_1
By verbal means	l_2	l_2
By physical and verbal means	l_3	l_3
Reward	g	*g*
Cooperation		
Casual or straightforward	+1	+1
Enthusiastic	+2	+2
Compliance	/	/
Noncooperation		
With explanation	−1	−1
With alternative	−2	−2
Simple refusal	−3	−3
Strong refusal	−4	−4
Ambivalent statement	u	*u*
Interpretation		*z*
Restatement of content	z_1	
Verbalization of motor behavior	z_2	
Restatement of verbalized feelings	z_3	
Recognition of feeling in motor behavior	z_4	
Clarification of verbalized feelings	z_5	
Clarification of feelings in total behavior	z_6	
Association of current and past events	z_7	
Account of reality or translation of symbolic behavior	z_8	
Anxiety		
Little or none	0	*0*
Some	+	+
Much	++	++
Hostility		
Little or none	0	*0*
Some	−	−
Much	−−	−−

* Underlined on scoring sheet as necessary.

General Recording Procedures

As soon as the observation begins, the behavior of adult and child is noted, identified in terms of categories, and recorded by placing the symbols for the categories in appropriate positions in the interval square. These positions are rigidly defined and adhered to throughout the recording. Regardless of initiator and of stimulus or response qualities, *the first behavior to appear within a 5-second period is always entered in the upper left corner of the time period square.* The location in the square of behavior subsequent to this initial entry depends upon whether adult or child provides the initial behavior. If the adult provides the initial behavior, the child's behavior is recorded in the lower right corner of the square. (Note: Symbols referring to child's behavior were underlined on the scoring sheets as necessary.) For example: The adult suggests the child build with blocks (d_1), and the child responds enthusiastically ($+2$). This recording is shown in Figure 1 (*a*), which also illustrates another recording principle, namely, that *the lower right corner of the square is used solely for recording the child's behavior.* If the child provides the initial behavior within a 5-second period, the adult's behavior is always recorded above the square. For example, when the child directs the adult (d_3), and the adult cooperates in a straightforward way ($+1$), the interaction is recorded as in Figure 1 (*b*). That is, *only the adult's behavior can be recorded in the space above the square.* The same scoring principles apply when the initiating behavior of adult or child is obviously not responded to, as when the child ignores (w_4) the adult's question (t_2), which is recorded as shown in Figure 1 (*c*).

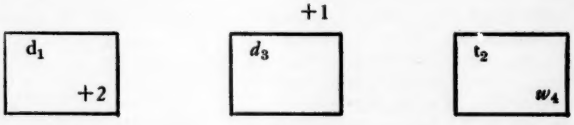

FIGURE 1

The same principles apply when both child and adult are not attentive (o), when both are attentive (c), or when one is attentive (c), while the other is not (o). The three possible combinations of these categories and their appropriate positions in the square are shown in Figure 2.

Since the time interval is short, the interaction in most 5-second periods can be accounted for with two entries, one for the child and one for the adult. Occasionally, however, it becomes necessary to make three recordings in one square. If the adult initiates the behavior in the record period, the third category entry will be the adult's and will appear above the square.

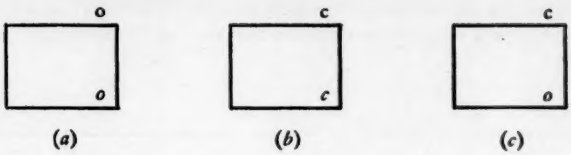

FIGURE 2

For example, the adult suggests that the child paint (d_1), the child asks about the brushes (t—), and the adult responds by telling the child where the brushes are (t_1). The recording of this behavior sequence is shown in Figure 3 (*a*). A similar procedure is followed when the child initiates the behavior. Figure 3 (*b*) shows the recording of a behavior sequence when the child asks a question (t—), the adult ignores it (w_4), and the child repeats the question (t—).

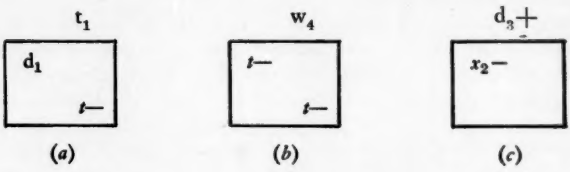

FIGURE 3

The previous illustrations have not included the anxiety-hostility rating, which actually accompanies each interactive category in the schedule. The symbol denoting the degree of anxiety or hostility immediately follows the behavior entry. For example, when the child, with an angry expression on his face, raises a hammer to strike the adult (x_2—), and the adult says, in an unsure voice, "Put the hammer down," (d_3+), the incident is recorded as shown in Figure 3 (*c*). Neutral, positive, or negative behavior expression in connection with nonattention (o or *o*) would be recorded as o_0, o_1, or o_2. For example, "child breaks car in solitary play" would be recorded as o_2.

When behavior extends beyond one 5-second period, it is recorded once every 5 seconds for as long as it continues. For example, when the child takes more than 5 seconds to ask a question (t—), and the mother listens during the first 5 seconds (c) and answers the question in the following

5-second period (t_1), a record is made as shown in Figure 4 (*a*). If, during observation of an interaction situation the child and adult separate and are out of visual range of each other, a dash is recorded in the time periods involved.

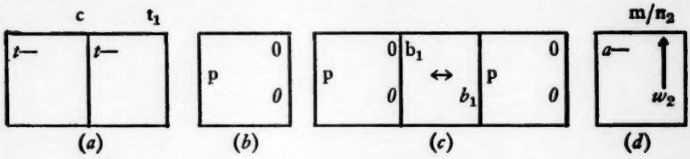

FIGURE 4

Where it adds meaning to the recorded symbols, a short descriptive account of the behavior of adult and child is placed on the scoring sheet, directly above the square or squares concerned.

Special Recording Procedures

Category p, Joint Participation in Activity, differs from all other categories in representing simultaneous interaction; that is, it does not separate the ongoing interaction into specific individual behavior. Consequently, only one category entry is used to indicate the behavior of both adult and child. In this instance only, the anxiety-hostility rating is recorded for the adult in the upper right corner of the square, and for the child in the lower right corner, as shown in Figure 4 (*b*).

When a specific interaction extends beyond a 5-second period, especially if it is Joint Participation in Activity or Giving Help, another interactive behavior often occurs simultaneously. For example, while adult and child are engaged in working a puzzle (p), they may carry on a discussion completely unrelated to the puzzle (b). In this case it is necessary to record two kinds of behavior in one 5-second period, as shown in Figure 4 (*c*).

When child or adult makes a statement that contains two independent clauses, which represent behavior fitting two different categories, each aspect of the statement must be recorded separately. Suppose, for example, the adult uses constructive criticism. The child shows the adult a drawing of a boy, which he has just completed (*a*—). The adult says: "That's the best drawing you have made (m), but perhaps if it had some arms it would look more like a boy" (n_2). The child may respond to the praise or to the criticism. Assuming that he responds to the criticism, saying, "It does too look like a boy!" (w_2), the sequence would be recorded as shown in Figure 4 (*d*), with an arrow connecting the response and its initiator.

A final point may be made in relation to the need for information indicating whether behavior occurs in response to an expressed need or some

other stimulus. Experience in using the schedule showed that such special subcategories were unnecessary, since the scores in the preceding time periods nearly always provide such information. For example, if Help (h) is given by the adult, and the symbol for Seeking Help, (*h*—), appears in the square immediately preceding, it would be evident that help was sought. If such a symbol does not appear, it may be assumed that help was given without being sought.

Observer Training

The observers' first task was to become thoroughly familiar with the details of the categorization system, their code letters, and the over-all scoring system. By extended discussion they arrived at common agreement concerning the specific types of behavior defined in the category system.

When the observers began to feel comfortable with the over-all details, they began observing the therapist and child from behind the screen. In the beginning they observed independently for one minute, then discussed each category entry that appeared within that time. An attempt was made to recall the specific kinds of behavior that gave rise to each recording, to analyze their disagreements, and to keep a record of the bases for them. Even when they agreed, the reasons for agreements were discussed, so that a clear understanding of the bases for scoring would result.

The bases for disagreements were then discussed among the authors of this paper. A further refinement of various categories was found necessary and was effected by more adequate definitions and division of categories, as a result of continuous interchange between authors and observers.

In view of the initial complexity of the task, the observers did not themselves keep time. A third person tapped out each passing time period and noted verbally a 30-second period, so that the observers could synchronize their observations. After the observers were familiar with the technique of recording, they began to practice timing themselves with a stop watch. Timing eventually became a rather simple task, for apparently a rhythm developed that made for accuracy. Nevertheless, periodic 30-second checks were made to assure synchronized recording.

After a month of intensive training, the observers were able to observe, more or less automatically, the complex parent-child or therapist-child interaction and enter the appropriate scores in the proper time blocks on the scoring sheet.

Reliability

Reliability of observations was established by obtaining observer agreement on identifying behavior in appropriate time sequence. The following formula (4) was used:

number of agreements / (number of agreements + number of disagreements)

In using a predetermined system of categories, in which the categories are recorded on a finely graduated time scale, several factors may operate to cause differences in the identification of behavior: the possibility of observers' seeing or hearing different events at the same time; imperfect familiarity with the categories; or an overly large number of categories in the schedule, making it impossible for any one observer to make reliable use of them.

It is sometimes impossible to determine whether disagreement between simultaneous records is a true difference in identification or only a timing error. When timing errors are possible, the computation of observer agreement becomes especially difficult. By inspection alone it is impossible to ascertain actual agreements and actual differences. To meet this situation we decided to establish certain arbitrary rules as to what constituted an agreement, a timing error, and an error attributable to other causes. When an identical behavior category or qualitative rating was recorded by each observer in the same time period, it was scored as agreement. A difference was scored as a timing error when the same category or rating was recorded by both observers in adjacent time periods. A difference that could not be attributed to timing error was considered a true disagreement.

Category Reliability

Category reliability refers to the ability of two observers to agree on identification of observed behavior. Individual category reliability measures for the three interactive situations appear in Tables 2 and 3. They are based on two 45-minute observation periods in each situation. Most of the categories show above 80 per cent agreement, indicating a reliability high enough to assure confidence in their use.

Only five categories had a reliability measure below 70 when computed on the total combined observations. Of these five, three were child categories (Q—, /, and w_1), and two were adult categories (z_3 and w_1). Each of these five categories had a total combined frequency of five or less and so required only one or two errors in agreement to yield low reliability. It is impossible to know whether these categories would prove reliable with increased frequency, but for practical purposes they can be accepted as functional.

Further observation is required to ascertain the specific reliability of categories that never appeared, the observers agreed, in the three situations observed. Such categories, quite possibly, would appear if observations were continued.

Observer Reliability

Observer reliability refers to the ability of two persons to agree upon the occurrence of an event in time. Contrary to the calculation of category reliability, timing errors are considered to be disagreements in calculating observer reliability. Only when an identical behavior category or qualitative

TABLE 2

RELIABILITY OF ADULT CATEGORIES

(Expressed in terms of frequency, and based on percentage of agreement for two 45-minute periods of observation in each of three adult-child interactive situations)

Category (Symbol)	T - C *Playroom*		M - C *Home*		M - C *Playroom*		*All Observations*
	fr.	%	*fr.*	%	*fr.*	%	%
o	3	25.00	397	89.92	30	100.00	91.51
c	1037	98.94	138	78.26	341	89.74	94.99
b	1	00.00	53	71.70	90	78.89	75.69
p	3	100.00	78	96.15	19	100.00	97.00
t_1	2	00.00	88	71.59	100	90.00	80.53
t_2	—	—	4	75.00	16	93.75	90.00
t_3	—	—	—	—	15	100.00	100.00
t_2—	3	00.00	104	75.96	73	95.89	82.78
h	19	78.95	28	96.43	4	100.00	90.20
s_2	5	100.00	—	—	—	—	100.00
s_{3a}	—	—	1	100.00	—	—	100.00
d_1	—	—	10	60.00	10	80.00	70.00
d_3	—	—	72	87.50	29	75.86	84.16
f_1	3	100.00	—	—	1	100.00	100.00
f_3	—	—	2	100.00	—	—	100.00
f_4	—	—	1	100.00	—	—	100.00
f_6	—	—	2	100.00	—	—	100.00
f_7	1	100.00	—	—	—	—	100.00
f_9	2	100.00	1	100.00	—	—	100.00
r_5	—	—	5	100.00	—	—	100.00
a	—	—	1	100.00	—	—	100.00
Q_2	—	—	1	100.00	1	100.00	100.00
Q—	—	—	1	100.00	—	—	100.00
m	—	—	7	85.71	4	100.00	90.91
n_1	—	—	2	100.00	—	—	100.00
l_2	—	—	1	100.00	—	—	100.00
l_3	—	—	10	90.00	2	100.00	91.67
z_1	15	60.00	11	72.73	7	100.00	72.73
z_2	6	100.00	—	—	1	100.00	100.00
z_3	7	28.57	—	—	1	100.00	37.50
z_4	11	90.90	—	—	1	100.00	91.67
z_5	5	100.00	—	—	—	—	100.00
+	—	—	3	100.00	2	50.00	80.00
−2	—	—	4	75.00	—	—	75.00
e	11	81.82	108	78.59	235	81.33	77.03
y	—	—	48	83.33	33	81.82	82.72
w_1	—	—	3	66.67	—	—	66.67

NOTE.—T = therapist; M = mother; C = child.

TABLE 3

RELIABILITY OF CHILD CATEGORIES

(Expressed in terms of frequency, and based on percentage of agreement for two 45-minute periods of observation in each of three adult-child interactive situations)

Category (Symbol)	T-C Playroom fr.	T-C Playroom %	M-C Home fr.	M-C Home %	M-C Playroom fr.	M-C Playroom %	All Observations %
o	1009	98.91	297	81.43	149	89.26	94.27
c	16	68.75	102	70.59	101	76.24	73.06
b	27	81.48	148	72.30	384	90.63	85.33
p	3	100.00	78	96.15	19	100.00	97.00
t_1	—	—	69	69.57	61	86.89	77.69
t_3	—	—	—	—	4	100.00	100.00
t_2—	—	—	58	77.59	64	87.50	82.79
d_3	9	77.78	20	85.00	15	100.00	88.64
s	2	100.00	1	100.00	4	100.00	100.00
h—	—	—	1	100.00	—	—	100.00
a	—	—	2	100.00	—	—	100.00
e—	—	—	31	83.87	—	—	83.87
Q_2	—	—	1	100.00	—	—	100.00
Q—	—	—	4	50.00	—	—	50.00
m—	—	—	—	—	3	100.00	100.00
l_3	—	—	1	100.00	—	—	100.00
l—	—	—	2	100.00	1	100.00	100.00
v_2	4	100.00	—	—	—	—	100.00
+1	1	100.00	25	80.00	31	80.65	80.70
+2	—	—	3	100.00	—	—	100.00
/	—	—	2	50.00	—	—	50.00
−2	—	—	2	100.00	—	—	100.00
e	25	76.00	112	79.46	71	76.06	77.88
w_4	29	75.86	56	82.14	37	78.38	79.51
w_1	—	—	—	00.00	—	—	00.00
w_3	—	—	—	100.00	2	100.00	100.00

NOTE.—T = therapist; M = mother; C = child.

rating is recorded by each observer in the same time period are the records considered to be in agreement.

Observer reliability was calculated for the categories, for the anxiety-hostility ratings, and for all entries within a time period, as shown in

Table 4. Observer reliability did not vary significantly in observations of mother and child in the home and in the playroom, or when these observations were compared with those of therapist and child. Reliability for the combined behavioral categories was 92.15; for the anxiety-hostility ratings, 99.10; and for the ratings within each time period, 88.51.

TABLE 4

OBSERVER RELIABILITY

(Expressed in terms of percentage of agreement on behavioral categories, anxiety-hostility ratings, and time period squares for two 45-minute periods of observation in each of three adult-child interactive situations)

	T-C *Playroom*	M-C *Home*	M-C *Playroom*	*All Observations*
Behavioral categories	96.41	87.58	92.72	92.15
Anxiety-hostility ratings	98.11	99.67	99.59	99.10
Time period squares	92.44	83.27	89.41	88.51

NOTE.—T = therapist; M = mother; C = child.

GENERAL TRENDS

This pilot investigation points to some tentative conclusions regarding adult-child interaction in the three types of situations studied.

In the frequencies of adult and child behavior (Tables 2 and 3), the behavior of adult and child tended to vary from one situation to another. However, there was greater similarity in the behavior of mother and child in the playroom and in the home than in either of these situations as compared with that of therapist and child in the playroom.

In the therapy situation, the range of behavior of both adult and child was comparatively limited. Approximately 93 per cent of the therapist's behavior consisted of attentively observing the child; that is, his most frequent response was in the nature of *being there,* watching and listening. Interpretation and Orientation were the only other approaches used to any considerable extent by the therapist.

The two children observed in play therapy had disturbing problems of severe anxiety and withdrawal. Their extreme inhibition in play necessarily limited the extent of interaction between child and therapist. These children were engrossed in solitary behavior approximately 95 per cent of the time. Nonrecognition or rejection of the therapist by ignoring or evading the therapist's stimulations was the most frequent interactive behavior, accounting for about 37 per cent of the child's total behavior. In frequency of occurrence, Statements of condition or action, Recognition of therapist's stimu-

lations, Attentive observation, and Directing the therapist by command, followed, each accounting for 2 per cent or less of the child's behavior.

In comparison, the home and playroom behavior of mother and child tended to be interactive in nature rather than solitary. In both home and playroom the mother's behavior was characterized by Attentive observation, Statements of condition or action, Offering verbal information, Seeking impersonal information, and Recognition of the child's stimulation. Directing the child by command was used often by the mother in the home, relatively infrequently in the playroom. Mother and child participated jointly in activities considerably more often in the home (78 episodes) than in the playroom (19 instances). While the mother showed 397 instances of Nonattention in the home, she failed to follow the child's behavior in the playroom in only 30 instances. The mother occasionally forbade, restricted, and criticized the child in the home, but only rarely in the playroom.

The children's behavior with their mothers in both home and playroom (Table 3) shows frequent instances of Statements of condition or action, Nonattention, Recognition of mother's stimulations, Attentive observation, Offering verbal information, and Seeking impersonal information. Rejection of the mother by ignoring or evading her stimulations occurred more frequently in the home (56 times) than in the playroom (37 times).

These trends are presented as tentative and suggestive patterns rather than conclusive ones, and require examination and checking against larger samples and more comprehensive data.

Discussion

The schedule presents general purpose categories covering a wide range of adult-child social interaction. With such a research tool, comparable studies of adult-child interaction can be made, offering, for example, an opportunity to compare parent-child and teacher-child interactions.[11]

The schedule intentionally allows an objective description of ongoing behavior, without assumptions concerning the motivational determinants of such behavior. However, a study of the types of behavior descriptions obtained from the schedule indicates that the behavior can be interpreted according to any theoretical position.

The schedule includes what the authors assume to be the more important categories of adult-child social interactional behavior. The question of how many categories an observer could learn and use effectively was the guide in determining how many categories to include. We believe the number

[11] Two studies, one published, one in progress, involve use of the schedule. The first, concerned with therapist-child interaction, has been published (10). The second, comparing mother-child interaction in playroom and home, is the doctoral dissertation of Henry D. Schalock, working under the direction of Dr. Sigel at The Merrill-Palmer School and a graduate committee at the University of Nebraska, Department of Education, headed by Dr. Warren R. Baller.

used in the schedule approaches the maximum for effective use; it includes, at any rate, the number that two observers could handle adequately.

Another problem concerns the reliability of all the categories. As shown in Tables 2 and 3, some categories never appeared in the sample observations. Conceivably, the appearance of some categories may require a very large sample. At this point we cannot delete the categories as useless, in view of the small sample, but must await a more extended use of the schedule.

The schedule has various limitations. Thus, in any preconceived category structure certain types of behavior will be omitted. In this schedule, for example, we are aware that much nonverbal social behavior, such as voice tone, gestures, postural signals, facial expressions, are omitted in favor of concentrating primarily on verbal behavior. Meaningful behavior of this kind might be the subject of further schedule construction. Another inherent limitation is that the schedule leaves the observer with a contentless description of the behavior observed. Verbal material is not available for the investigation of such questions as the symbolic significance of the response, language development or usage, and the like. Some other well-known limitations of such an observational procedure are problems of time sampling, the shortcomings of preconceived observation schedules, and the like.

Limitations notwithstanding, we offer this technique as useful for the observational purposes outlined. Further research should refine the technique further, and so increase its usefulness and extend its applicability.

Summary

A schedule is presented for observation of adult-child interaction, with a total of 89 adult and 82 child categories, as well as anxiety-hostility ratings. The schedule has been tested and found reliable in studying parent-child interaction in home and laboratory (playroom) and in studying therapist-child interaction. Certain general trends, similarities, and differences in mother-child interaction in home and playroom and in therapist-child interaction in the playroom are noted. The general applicability, strengths, and weaknesses of the procedure are discussed.

References

1. Arrington, Ruth E. Time sampling in studies of social behavior: A critical review of techniques and results with research suggestions. *Psychol. Bull.*, 1943, 40, 81-124.

2. Bishop, Barbara Merrill. Mother-child interaction and the social behavior of children. *Psychol. Monogr.*, 1951, 65, No. 11.

3. Collier, R. M. A scale for rating the responses of the psychotherapist. *J. consult. Psychol.*, 1953, 17, 321-326.

4. Festinger, L. *Analysis of methods of calculation of per cent agreement.* Laboratory Course in Child Psychology, 1944-45. (Mimeographed)

5. Finke, Helene. Changes in the expression of emotionalized attitudes in six cases of play therapy. Unpublished master's thesis, Univer. of Chicago, 1947.

6. Landisberg, Selma, & Snyder, W. U. Nondirective play therapy. *J. clin. Psychol.*, 1946, 2, 203-214.

7. Lebo, D. The relationship of response categories in play therapy to chronological age. *J. child Psychiat.*, 1952, 2, 330-336.

8. Moustakas, C. E. *Children in play therapy.* New York: McGraw-Hill, 1953.

9. Moustakas, C. E. The frequency and intensity of negative attitudes expressed in play therapy: A comparison of well-adjusted and disturbed young children. *J. genet. Psychol.*, 1955, 86, 309-325.

10. Moustakas, C. E., & Schalock, H. D. An analysis of therapist-child interaction in play therapy. *Child Develpm.*, 1955, 26, 143-157.

11. Osborn, D. K. Dependency in young children in relation to two maternal variables: A methodological study. Unpublished master's thesis, State Univer. of Iowa, 1952.

12. Seeman, J., *et al.* The development of empirical criterion measures in the study of play therapy. (Unpublished manuscript)

IDENTIFICATION WITH PARENTS AS A DETERMINANT OF DOLL PLAY AGGRESSION[1]

HARRY LEVIN and ROBERT R. SEARS
Cornell University *Stanford University*

A major problem in the study of personality development is to discover behavior variables that not only are relevant to the more important motivational systems of the child but which also permit precise measurement and some degree of experimental control. The so-called projective techniques have seemed promising in these respects. They provide responses belonging to significant motivational categories such as sex, aggression, dependency and achievement, but the responses are highly attenuated in their intensity and may be elicited safely and conveniently under controlled conditions. Projective responses have no special virtues other than these; they are not, in themselves, "true" measures of any "deep motives," nor do they have any unique validity as measures of personality qualities. They are simply acts that have presumably discoverable antecedents and that bear regular relationships to other acts. Fantasy behavior, in other words, is no different from any other behavior in its logical status as a reference event to be accounted for by principles of learning and action.

Experimental doll play has been rather widely used for the study of fantasy aggression in preschool aged children, both because this method requires less verbal skill than the visual-perceptual projective techniques and because it seems to be peculiarly effective in eliciting aggressive activity. Doll play offers an opportunity for measurement of three aspects of fantasy

[1] This article reports a selected set of findings from a larger study of "Identification in Young Children" conducted by the Laboratory of Human Development at Harvard University. Primary responsibility for the study rested with Robert R. Sears, Harry Levin, Eleanor E. Maccoby, Pauline S. Sears, and John W. M. Whiting. The study was made possible by a grant from the National Institute of Mental Health (M-461). Parts of this paper were presented at the 1953 meetings of the American Psychological Association in Cleveland, Ohio.

aggression: (*a*) the *amount,* as indicated either by frequency of aggressive behavior units or by proportion of time devoted to aggression, (*b*) the *agents and objects* of aggression, i.e., who attacks whom, in thematic play with doll characters, and (*c*) the *quality* of the aggression in terms of such categories as directness, intensity, verbal or physical expression, and pro- or contra-social quality. These various attributes of aggression can be measured by the use of appropriate recording categories, and the resulting quantitative "scores" can then be used as standard reference events for which correlates can be sought. While these fantasy events are of considerable interest in their own right, they have added importance that comes from their, as yet, largely undiscovered relationships to comparable aspects of overt social aggressive behavior, and their inhibiting or facilitating influence on non-aggressive activity.

Previous researches have discovered some of the factors that influence these doll play measures. We will limit our consideration here to those that have been found to be related to the *amount* of aggression, since this is the measure with which the present study is concerned.

There are several aspects of the methods used for doll play that are relevant (10). Among these are (*a*) the amount and kind of interaction between the child and the experimenter, with a high level of interaction initiated by the latter being positively related to frequency of aggressive response; (*b*) the realistic quality of the doll house materials, the more realistic being favorable to aggression; (*c*) the duration of the play session, with a 20-minute session producing a relatively greater amount of thematic aggression than a full hour session, which tends to elicit a good deal of tangential behavior; and (*d*) degree of organization of the doll house and furniture, the more organized setup providing higher average aggression scores.

The actual handling of the child during his play session is also important. Experimental doll play is customarily done under conditions of high permissiveness for aggression, the experimenter adopting an interested but neutral attitude toward the child's behavior. With this procedure there is usually an increase in frequency of aggressive acts from the beginning of a session to its end, and also from session to session. When this permissiveness is replaced by what is perhaps a socially more customary attitude, one that involves reproof or disapproval for thematic aggression, the frequency of this kind of act decreases at once (3). It appears that the permissive attitude is anxiety reducing: it eliminates the constraint that the usual social control by adults ordinarily imposes on the child's aggressive activity, whether reality-oriented or fantasy. In sum, these methodological factors provide variations in the immediate instigating conditions.

There are other sources of influence, however, that are not subject to the experimenter's control. These are the child's relatively stable response potentialities that have developed either from his past experiences or are

characteristic of his constitution. Some children are more aggressive than others in the doll play situation. In general, boys have been found to be more aggressive than girls (7, 8), accident-prone boys more so than non-accident-prone ones (5), and boys with their fathers living at home more so than those with their fathers absent from home (8, 12).

While it is probable that constitutional factors have some influence on these aggressive response potentials, our present concern is with the effects of differences in child-rearing experiences. The studies just mentioned are of some value in this respect, for they refer to groups which have been or could be examined with respect to the kinds of child training used by the children's mothers. For example, as has been shown elsewhere (11), there are important differences in the treatment accorded boys' and girls' aggression in early life, and these differences are relevant to the sex differences in expression of fantasy aggression.

There are many other differences between boys and girls, however, and it is difficult to assess the exact influence of all of them on doll play aggression. A more direct approach can be made by holding constant such gross sources of influence as sex of child, and searching for the effects of specific child-rearing practices as these are revealed in interviews with the mothers of children whose fantasy aggression is measured. This was the procedure used in an earlier investigation (13), which provided a crude assessment of the effects of one child-rearing variable on amount of fantasy aggression, and suggested the hypotheses that have been tested in the present research.

Previous Findings

In the earlier study (13), 40 mothers of preschool aged children were interviewed about their child-rearing methods. From these reports, the mothers were rated on a scale that described the dimension *severity of punishment for the expression of aggression*. Their children were tested with two 20-minute sessions of doll play. Severity of punishment proved to be positively related to the amount (frequency) of doll play aggression. For the girls, the correlations were +.41 and +.46 for the first and second sessions respectively; for the boys, +.46 and +.23.

It is evident, of course, that this one child-training dimension was not accounting for a very substantial proportion of the total variance of the fantasy aggression scores. Another finding from the same study drew attention to the possible influence of *identification with parents*. The children were also observed in preschool, and measures were obtained of the amount of overt social aggression exhibited there. For the boys, the correlation of this measure with severity of punishment was +.56. But for the girls, the relation was curvilinear, with both high and low punishment associated with low overt aggression. It was as if the girls had been more severely punished on the average, although the ratings did not show this, and the

highest punished ones had become inhibited and passive in their social expression of aggression. The hypothesis was suggested that girls, being more highly identified with their mothers, suffered more inhibitory effects from a given degree of severity of punishment.

Hypotheses about Identification

Our attention having been brought to the identification process, we considered what possible effects it might have on expressions of fantasy aggression. Its relevance seemed to lie, possibly, in the character of doll play as a situation that permits a child to play the role of whichever family-member he wishes. There is a presumption that he identifies in some degree with all the persons who both reward him and place demands on him; this includes all the members of his family, but preeminently the parents. In doll play he can perform these identified-with roles, and do the kinds of things that he understands those roles require. Thus, the higher the identification a child has with a parent, the more his doll-play behavior should be like the behavior that the child believes is characteristic of that parent.

In the present context, we will assume that aggressiveness is a quality in a parent's behavior that a child can recognize, and that the child can discriminate whether the parent shows much or little aggression in relation to other (nonaggressive) forms of action. Thus, if we also assume, as suggested above, that in doll play a child adopts the role of his most-identified-with parent and acts out that parent's characteristics, we arrive at our first hypothesis:

(*a*) The more strongly a child is identified with a given parent, the more nearly will he approximate in doll play the level of aggression he perceives as characterizing that parent. This means that if a boy is strongly identified with his father, and the father is very aggressive, the boy will display a high level of aggression in doll play. Conversely, if the father is relatively nonaggressive, the boy will be also. If, however, the boy's identification is weak, the father's level of aggressiveness should have little or no influence.

But with which parent does a child identify? By the age of five years, most children may be assumed to have developed their strongest identification with the same-sexed parent. This development is not the same for the two sexes, because the conditions for establishing primary identification are such that the mother is the initial identificand for both boys and girls. Only as the boy begins to perceive himself as a male, "like" his father and "unlike" his mother, can he begin to shift his identificand from mother to father. Hence, identification is more straightforward in the girl, and she develops a stable sex-typing more rapidly in this respect than the boy.

This latter assertion is supported by two findings from previous studies of doll play. If we accept the assumption that the doll most frequently chosen as agent of an act is the fantasy representative of the chief identificand, we can examine frequency of choice by children of both sexes. P. S. Sears reported (9) that the five-year-old girls of the present study used the mother doll most frequently for neutral positive (nonaggressive) acts, whereas boys used the mother and father dolls almost equally in the first session and used the father more only in the second session.

Another type of comparison was made by Johnson (4). She compared the relative frequency of pro-social and contra-social aggressive acts in five- and eight-year-old children of both sexes. The pro-social acts were chiefly disciplinary ones that represented *adult quality* of behavior. Both sexes increased sharply in this type of aggression from five to eight years of age, but the girls were very much higher than the boys at both ages, and the five-year-old girls were almost as high as the eight-year-old boys.

We can conclude from these findings that the chief identificand for a child at five to six years of age is most commonly the parent of the same sex, though the relative strength of such an object choice for identificand is greater for girls than for boys. This leads to our second hypothesis:

(*b*) The father's level of aggression will be most influential in determining the boy's frequency of aggressive doll acts, and the mother's level will be most influential for the girl's.

But what determines the child's perception of his same-sexed parent's aggressiveness? There appear to be two types of experience that are important, and their interaction produces a curious and complex difference between the sexes with respect to what child-training variables influence the frequency of doll play aggression.

The first of these two factors is perception of sex role. Characteristically the male is more aggressive. Among adults, there are more suicides, homicides and street fights by men. War has always been a male preoccupation. Rough contact games are mainly for males. Even in nursery school the sex difference is evident, in both overt and fantasy expression. The boy may be expected to perceive his father as a representative of males in general, and hence as relatively aggressive.

The other factor is the child's own *personal experience of parental aggression*. Even though a boy recognizes the aggressiveness of the male role, his own father may be relatively mild. The opposite variation can occur for a girl. For most children, it seems likely, the clearest measure of parental aggressiveness is derived from the severity with which the parents punish the child for his aggressive actions toward them. If his aggressive acts are met with vigorous counter-aggression—i.e., parental actions that the parents themselves view as punishment—the child will perceive parents as more aggressive than if his acts are met with only mild punishment. The important aspect of this experience lies in the distinctiveness of the cues it

provides the child. Presumably, when the chief punisher is the parent with whom the child identifies more strongly, the cues will be more distinctive. Hence we get the following hypotheses:

(*c*) The frequency of doll play aggression will be greater in boys than in girls (because of the male role perception). This hypothesis has been verified by a number of studies (1, 7, 8).

(*d*) The frequency will be greater when punishment for aggression is severe than when it is mild. This was found to be the case in one previous study, mentioned earlier (13).

(*e*) Given severe punishment for aggression, the frequency of doll play aggression will be greater when the parent of the same sex is the chief disciplinarian.

If we now convert these five hypotheses to statements of what child-training antecedents will contribute to the frequency of children's doll play aggression, we find that the factors are different for the two sexes. For boys, we would expect the following for the reasons indicated: (*a*) *degree of identification* with parents, because the male role is an aggressive one, (*b*) *severity of punishment,* because this adds information for the boy's perception of his identificand's level of aggressiveness, and (*c*) *the father is the chief disciplinarian,* because this increases the distinctiveness of cues and spells out unmistakably to the boy that males are aggressive. In all likelihood, too, this increases the severity of frustration the boy suffers, because the punishment comes from the identificand, and hence adds to the boy's motivation for aggressive counter-actions.

For girls, the factors are different. Since the female role is *not* an aggressive one, the girl's identification with her mother should not in itself be a significant antecedent except as it may be associated with a mother who is a very aggressive person.

The antecedents should be (*a*) *degree of identification,* but only when there is high punishment, that is, the girl is provided with an atypical female model, (*b*) *severity of punishment,* because this provides an aggressive role model, but only when (*c*) *the mother is the chief disciplinarian.*

Method

The testing of these hypotheses required measurement of four variables. One was the criterion, or consequent, variable of frequency of aggressive acts in the doll play of five-year-old children of both sexes. The second was the degree of identification of these children with their parents. The third was the severity of punishment for aggression that the parents used on the children, and the fourth, an indication of which parent was the chief punisher of the child.

In connection with an extended study of the origins of identification in a group of 379 five-year-old children, these four measures were obtained. The last three were secured through interviews with the children's mothers.

The children were all enrolled in public school kindergartens in two suburbs of a New England metropolitan area. There were 202 boys and 177 girls, of various ordinal positions, and the families were fairly evenly spread from upper lower socio-economic status to upper middle. The mean age of the subjects was 5-7, the range being 4-11 to 6-7, with a standard deviation of 4.4 months. The details of this sample have been described elsewhere by Maccoby and Gibbs (6).

Doll Play

Each child was given two 20-minute sessions of doll play, never on consecutive days, but always within a five-day period. The play was scored in terms of behavior units with but two categories, aggression and neutral positive acts. Aggression was defined as an action that had the intent to irritate, hurt, injure, punish, frustrate or destroy a doll or equipment. It ranged in intensity from innocuous mischief (the girl sneaks out of bed) to severe accidents, fights and catastrophes (the father spanks the boy, the father has half his head sliced off, a bomb falls on the house and kills everybody). Verbal aggression such as scolding, threats, and uncooperativeness were included, as were instances in which discomfort was attributed to a doll (the boy is sick, sad, lost, etc.). The doll agent of each act was also recorded. The measure used in the present study is the "per cent aggression" for each session separately. This value is obtained by dividing the frequency of aggressive behavior units in a session by the total number of units of play. Nine experimenters collected these data. They had been trained to an adequate degree of comparability in their scoring, so their data have been combined. A full account of this doll play procedure is presented elsewhere (8).

Mother Interviews

The children's mothers were interviewed in their own homes by women interviewers who had been trained to use a standardized set of questions that covered the major areas of child rearing. This interview, which has been described elsewhere (11), consisted of 72 main questions and appropriate probe questions to assure that the necessary information would be obtained on each topic. The interviews were recorded and then were transcribed to typescript. They averaged between two and two and a half hours in length, and approximately 50 pages in single-spaced typed form. A series of rating scales and category codes was used for quantifying the interview information. Each interview was rated on every scale by two raters working independently. The reliabilities of the ratings were calculated by correlations between these independent judgments. The final score used for each scale on an interview was the pooled rating of the two raters; when the two ratings were seriously discrepant, the judges reexamined the interview and attempted to reach a closer agreement. Reliability coefficients were calculated *before* any such discussions.

From the interviews, measures were obtained of the three variables that were required for the present analysis. These, together with their Spearman-Brown corrected reliability coefficients, were (*a*) degree of identification with parental role that was displayed by the child ($r = .77$), (*b*) the severity of punishment given the child for expressing aggression toward the parents ($r = .77$), and (*c*) which parent was the usual punisher ($r = .95$). Some comments are needed to clarify the bases on which ratings on these three scales were made.

Identification

The rating of the child's degree of identification was based on the mother's replies to the following questions:

1. We'd like to get some idea of how X acts when he's naughty . . . when he has deliberately done something he knows you don't want him to do, when your back is turned; how does he act?
2. Does he ever come and tell you about it without your having to ask him?
3. When you ask him about something he has done that he knows he's not supposed to do, does he usually admit it or deny it?
4. What do you do about it if he denies something you are pretty sure he has done?

We reasoned that the identification process is necessary to the formation of super-ego, or conscience, and that the amount of identification the child has with parental roles can be estimated from the stage of development he has reached in internalized control. These four interview questions were designed to bring out the mother's observations of her child's behavior that were relevant to this matter.

The rating scale was labelled *evidence of super-ego in child,* and the raters were instructed to take into account the following: "High super-ego is inferred from confession of deviation, acting guilty and feeling bad over deviations, admitting rather than denying a fault. Hiding and other evidences of fear should be taken as evidence of *low* super-ego. The making of reparations is evidence of *high* super-ego. If any of the indicators of high super-ego are explicitly stated to have occurred because of fear, however, they should be discounted. For example, if child admits rather than denies because he has been severely punished for denying, admitting is less direct evidence of super-ego."

The following descriptive statements were used to define the five points on the rating scale; the per cent of children rated at each point is also given.[2]

[2] The two raters disagreed by one point, on this five-point scale, on 109 children. For purposes of this tabulation, half of the disagreed-upon cases were allocated to each of the adjoining scale points. For example, there were 35 children that both raters agreed should be "1" and 72 that both agreed should be "2." But there were 28 cases that were rated "1" by one rater and "2" by the other. Half of these, or 14 have been allocated to "1" and half to "2." For all data analyses, those cases have been defined as "1.5."

1. No evidence of super-ego. Child denies, hides, doesn't seem unhappy when naughty.	13%
2. Slight super-ego.	28
3. Moderate super-ego. May not confess directly, but looks sheepish; seldom denies.	38
4. Considerable super-ego. May come and express affection to mother when he has been naughty; tries to repair damages.	17
5. High super-ego. Feels miserable when naughty; always confesses, never denies; strong need for forgiveness.	3
Not ascertained.	1
Total	100%

The following example will indicate the type of report that was rated toward the high end of the super-ego scale:

I: We'd like to get some idea of how Ed acts when he's naughty; when he deliberately does something he knows you don't want him to do when your back is turned, how does he act?

S: Very seldom does that. But a few times he has done something that he shouldn't do, that I don't know anything about, if I'm in the other room, he just can't hold it in very long and finally comes in to me and he says, "Mother," and I'll say, "What?" "I did something I shouldn't have done." Instead of leaving it and getting away with it he usually comes over and tells me what he's done. He usually comes, I mean, and it's not very long after he's done it. He can hardly hold it in to himself, you see.

One anecdote that was volunteered by a mother, although not necessarily an example of high super-ego, suggests the way in which a child can display adult attitudes and play the adult role; it shows how a highly identified child takes over some normally parental behavior:

S: I've taught him never to touch a medicine cabinet. I showed my little boy bottles—medicine bottles—and I've told him—he was sick, oh, about a year ago, in the hospital; and he's never forgotten it, and I don't think he ever will. He's told his brothers about it plenty of times. And I told him, "If you took your medicine bottle, you're going to get sick like you did then." He won't go near that medicine cabinet. And if he sees one of his brothers going to touch it, he'll tell him, "You'll go to the hospital, and you'll stay in the hospital for a long, long time; and you won't see Mommie or Daddy. And you won't have no toys to play with—just lay in bed all the time and get needles." And now none of them will go up there. I think he's really the one that broke them of going to the medicine cabinet.

Severity of Punishment for Aggression Toward Parents

An extended report of this scale and of the influence of such punishment on overt aggressive behavior in the home is published elsewhere (11). The scale was a five-point one that took account of how severe the punish-

ment was that the child actually received, regardless of who did the punishing.

Which Parent Was the Usual Punisher

The mother was asked, "Who usually disciplines, you or your husband?" The replies were categorized as "husband," "wife," or "about equal." Of the 379 cases, 108 were judged to fall in this last category, and 11 could not be evaluated. Since the hypotheses with which we are concerned here relate only to the condition in which the preponderance of punishment is by one parent or the other, we have limited the data analysis to the 260 children whose mothers reported either "husband" or "wife" as the usual punisher.

Since there were a few cases which could not be evaluated on the *identification* and *severity of punishment* scales, the total number of children actually utilized for the present analysis was 241, of whom 126 were boys and 115 were girls.

Test of the Hypotheses

The various hypotheses outlined above offered certain predictions as to what combinations of the three mother interview variables would be associated with high or low frequency of doll play aggressive acts. These predictions have been tested by a triple classification analysis of variance, separately for the sexes. Except for *who usually punishes,* which contained two mutually exclusive categories, the variables were continua; it was necessary, therefore, to break each distribution into as nearly equal halves as possible for the analysis of variance. It was impossible to secure exactly equal numbers of cases in the eight required cells, however, and hence the correction for unequal cells as given by Walker and Lev (14) was used. The analysis was done separately for each of the two doll play sessions.

Boys

The analysis for boys is shown in Table 1, which gives the mean % Aggression scores for the eight groups of children representing the eight possible combinations of the three antecedent variables. The table also shows that in both sessions there were two major sources of variation in the % Aggression scores. One of these was *identification* and the other an interaction between *identification* and *agent of punishment.* The influence of these sources of variation is clearly shown in the means; the two highest means for the first session are "Father punishes, High Identification." On the second session, two of the three highest were in the same category.

The effects of these two variables can be more easily seen when the data are regrouped, as in Table 2. When the mother punished, the degree of identification had no influence on the amount of aggression displayed, but when the father punished, high identification produced high aggression.

TABLE 1

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS, AND ANALYSIS OF VARIANCE ACCORDING TO IDENTIFICATION, SEVERITY OF PUNISHMENT FOR AGGRESSION, AND AGENT OF PUNISHMENT (BOYS ONLY)

	High Identification				Low Identification			
	Severe Punishment		Mild Punishment		Severe Punishment		Mild Punishment	
	Father	*Mother*	*Father*	*Mother*	*Father*	*Mother*	*Father*	*Mother*
Session I	18.6	13.9	23.0	13.8	5.8	12.1	10.2	13.6
Session II	28.8	20.1	25.9	26.7	12.4	23.9	18.7	23.4
N	18	11	18	13	14	18	14	20

	Session I				Session II			
Source	*df*	*Sum of Squares*	*F*	*p*	*df*	*Sum of Squares*	*F*	*p*
Agent of Punishment	1	2.21	.19	n.s.	1	8.61	.52	n.s.
Severity of Punishment	1	13.00	1.10	n.s.	1	11.28	.68	n.s.
Identification	1	95.22	8.08	<.01	1	66.70	4.04	<.05
Agent × Severity of Punishment	1	6.85	.58	n.s.	1	.91	.05	n.s.
Agent × Identification	1	69.62	5.91	<.05	1	72.60	4.39	<.05
Severity × Identification	1	.32	.03	n.s.	1	.55	.03	n.s.
Agent × Severity × Identification	1	.32	.03	n.s.	1	33.21	2.01	n.s.
Residual	118	1389.84			118	1950.91		

These findings confirm the hypotheses that place emphasis on the boys' identification with an aggressive male role model as a determinant of the frequency of aggression in doll play.

One hypothesis is not supported, however. The severity of punishment for aggression proved to have no influence on doll play aggression. This finding is in contrast to that from the earlier investigation, although in that one the correlation of this variable with second session doll play

TABLE 2

MEAN % AGGRESSION IN TWO DOLL PLAY SESSIONS ACCORDING TO THE DEGREE OF IDENTIFICATION AND THE AGENT OF PUNISHMENT (BOYS ONLY)

	High Identification		Low Identification	
	Father Punishes	*Mother Punishes*	*Father Punishes*	*Mother Punishes*
Session I ...	20.8	13.8	8.0	12.9
Session II ..	27.3	23.7	15.6	23.6
N	36	24	28	38

aggression was quite small in boys (.23), while the girls' was larger (.46). Since neither *identification* nor *agent of punishment* was measured in that study, it is possible that the small positive correlation was a product of some unknown correlation with these variables.

We have no certain explanation for the lack of influence of severity of punishment, on boys, in the present study. One possible reason may be that, by the age of five years, most boys have been so much exposed to aggressive male models that the added experience of observing the father's being aggressive has little or no added effect. Possibly boys get enough training in male role aggressiveness that the only factors required for making them behave aggressively themselves are a strong identification with the adult role and the clear cues that derive from paternal punishment.

TABLE 3

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS, AND ANALYSIS OF VARIANCE ACCORDING TO IDENTIFICATION, SEVERITY OF PUNISHMENT FOR AGGRESSION, AND AGENT OF PUNISHMENT (GIRLS ONLY)

	High Identification				Low Identification			
	Severe Punishment		Mild Punishment		Severe Punishment		Mild Punishment	
	Father	*Mother*	*Father*	*Mother*	*Father*	*Mother*	*Father*	*Mother*
Session I	8.8	13.2	11.0	5.7	5.2	9.5	8.1	7.8
Session II	15.4	20.1	12.3	8.5	7.7	11.3	13.6	10.4
N	9	18	9	24	11	13	11	20

	Session I				Session II			
Source	*df*	*Sum of Squares*	*F*	*p*	*df*	*Sum of Squares*	*F*	*p*
Agent of Punishment	1	1.20	.16	n.s.	1	.21	.10	n.s.
Severity of Punishment	1	2.10	.28	n.s.	1	11.76	1.06	n.s.
Identification	1	8.20	1.20	n.s.	1	22.11	1.99	n.s.
Agent × Severity of Punishment	1	25.56	3.42	<.07	1	29.26	2.64	<.10
Agent × Identification	1	3.00	.40	n.s.	1	.03	.00	n.s.
Severity × Identification	1	5.28	.71	n.s.	1	48.51	4.37	<.05
Agent × Severity × Identification	1	3.25	.43	n.s.	1	.36	.03	n.s.
Residual	107	800.30			107	1187.90		

Girls

The case is quite different with girls. As Table 3 shows, the interaction between *severity of punishment* and the *mother-as-agent* of punishment approaches statistical significance in both doll play sessions. The data are rearranged in Table 4 to show the influence of these two variables. When the mother was the punisher and the punishment was severe, the fantasy aggression was greater than under any other combination of variables.

TABLE 4

MEAN % AGGRESSION IN TWO DOLL PLAY SESSIONS ACCORDING TO THE SEVERITY OF PUNISHMENT FOR AGGRESSION AND AGENT OF PUNISHMENT (GIRLS ONLY)

	Severe Punishment		Mild Punishment	
	Father Punishes	*Mother Punishes*	*Father Punishes*	*Mother Punishes*
Session I ...	6.8	11.7	9.4	6.6
Session II ..	11.2	16.4	13.0	9.4
N	20	31	20	44

This finding supports the prediction that exposing girls to an aggressive role model increases their doll play aggression. It is to be noted that the converse holds true as well; when the mother was the punisher, but punishment was mild, doll play aggression was least. In other words, the mothers who did not punish severely permitted their daughters to take over the socially modal nonaggressive female pattern. Severely punitive mothers erected potential barriers against their daughters' copying of at least this one aspect of approved femininity.

The role of identification appears to have been of some importance in this connection. In the second doll play session, there was a significant interaction between identification and severity of punishment. The data are regrouped in Table 5 to show the joint effects of these two predictor variables. Girls who were both highly identified and severely punished exhibited more aggression than the girls falling in the other three combinations. These latter did not differ from one another, on the average, since the remaining influential variable (agent of punishment) varied randomly among the groups.

This finding adds further support to the prediction that identification will lead to much aggression for girls only when it is associated with an atypically aggressive female role model, as exemplified in a severely punishing mother.

TABLE 5

MEAN % AGGRESSION IN THE SECOND DOLL PLAY SESSION ACCORDING TO THE DEGREE OF IDENTIFICATION AND THE SEVERITY OF PUNISHMENT FOR AGGRESSION (GIRLS ONLY)

	High Identification		Low Identification	
	Severe Punishment	*Mild Punishment*	*Severe Punishment*	*Mild Punishment*
Session II ..	18.5	9.5	9.6	11.5
N	27	33	24	31

Comparison of the Sexes

There are two striking sex differences in the antecedents of doll play aggression frequency. One is with respect to identification and the other to severity of punishment. In boys, identification is a significant positive influence, both alone and in interaction with other variables. In girls, it is apparently less so. There are two possible explanatory factors to be considered. One is that the measure of identification is of less discriminating value with girls than with boys, because girls are developmentally farther along in the process. As was mentioned earlier, boys presumably develop identification with their mothers first and then must shift to fathers for the appropriate role model. Girls, in other words, may have reached an asymptotic level of identification in their sixth year, while the boys of this age are strung out along a fairly wide range of identification. Thus this variable would assume greater significance for boys.

The other factor has to do with severity of punishment. The male role is an aggressive one, and the actual day-to-day demonstration of this fact by the father's use of severe punishment may add little or nothing to the boy's understanding of how he should behave. It is more important for the boy that he have high identification with the role. For girls, however, the appropriate sex role model is normally nonaggressive. Only when the particular model most visible to the girl (her mother) is especially aggressive does the girl adopt an aggressive way of behaving.

There is one further hypothesis we might suggest for further research on this problem of sex differences. Among the children in this sample, the boys were subjected to more *physical punishment,* according to the mothers' reports (11). If the aggressive behavior in doll play were scored as to whether it was physical or verbal in content, we would therefore expect that the boys would show more physical and the girls more verbal expressions. This sex difference was found in studies by P. S. Sears (8) and by Bach (1), but no child training antecedent measures were available. In the present study, the differentiation of content was not scored.

Identification with the Aggressor

Extended clinical study of children led Anna Freud to the discovery of a defense mechanism she has termed identification with the aggressor (2). This mechanism is a technique by which a child can reduce his fear of an aggressor, perhaps a punitive parent, by identifying with that person and adopting his aggressive behavior.

The findings of the present study are fairly well in accord with what one would predict from this mechanism if one assumes that frequency of doll play aggression is a measure of the aggressive motivation derived from that identification process. It is the girls who have most to fear from their mothers, i.e., the daughters of severe punishers, who show the highest aggression. The same does not hold for the boys, but for both sexes it is

the children who are ordinarily punished by the *parent of the same sex* who show greatest aggression. The connection between the measure of identification and amount of aggression may equally well be interpreted either as evidence that identification with the aggressor had taken place or, as we have done above, as an indication of the requirement of identification in order that agent of punishment be effective. In other words, the rating of high identification in those children (especially boys) who showed high aggression could be viewed as an *effect* as well as a *cause*.

We are inclined to rely on the interpretation we presented in the introduction to this paper rather than on the conception of identification with aggressor because we do not have any way of determining, from the present data, whether the fantasy aggressions had anxiety-reducing effect. This is an essential assumption underlying an interpretation in terms of this defense mechanism. As the data stand, however, there is no factual reason why the doll play aggressive acts cannot be interpreted as evidence of the operation of this mechanism.

TABLE 6

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS, WITH CORRELATIONS AND DIFFERENCES BETWEEN SESSIONS AND SEXES

	N	Session I	Session II	*r*	*t*
Boys	202	14.28	21.84	.61**	3.30**
Girls	177	8.64	12.76	.69**	2.25**
t between sexes		4.32**	4.95**		

** Significant at $p < .01$.

NORMATIVE FINDINGS

Since standardized doll play as used in this study is a common method of research with children, it is useful to catalog norms of such behavior. An analysis of the influence of sex, age, ordinal position and father's absence on doll play aggression has been reported by P. S. Sears (8). The present results, where applicable, will be compared with hers.

Sex Differences

As has been stated earlier, the boys devoted significantly more of their actions to aggression in both doll play sessions than did the girls (Table 6). This confirms other findings (1, 4, 7, 8) and may be taken as a reasonably stable characteristic of doll play behavior.

Session Changes

For both sexes, there was a significant session-to-session increase in the proportion of aggression units (Table 6). This finding is also in line with

previous studies (3, 8) and has been interpreted to mean that the deliberate permissiveness of the experimental procedure progressively acts to reduce the inhibitions about aggression that the child has heretofore acquired.

Although there was a session-to-session increase in the mean % Aggression scores, there was a correlation between sessions of .61 for boys and .69 for girls. P. S. Sears (8, p. 35) reported session-to-session correlations of .74 for five-year-old boys and .53 for five-year-old girls, with 25 cases in each category.

Ordinal Position

P. S. Sears (8, p. 24) found that, for both sexes, children who were "youngest" in their family or who were "only" children showed more aggression than "oldest" children. In the present study, four ordinal positions were analyzed: (*a*) *only* children, (*b*) *oldest* of two or more children, (*c*) *youngest* of two or more children, and (*d*) *middle* of three or more children. The results are presented in Table 7.

TABLE 7

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS ACCORDING TO THE SEX AND ORDINAL POSITION OF SUBJECTS

	BOYS			GIRLS		
	N	Session I	Session II	*N*	Session I	Session II
Only	33	19.36	30.64	19	8.88	13.76
Oldest	59	13.08	19.32	50	8.76	12.84
Youngest	62	14.68	18.72	51	8.48	13.36
Middle	48	13.28	22.84	57	7.84	11.76

The present findings only partially replicate the earlier ones. In both doll play sessions, "only" boys exhibited the most fantasy aggression, but among the girls, there are no sizeable mean differences in % Aggression which are attributable to ordinal position.

Descriptive categories such as ordinal position or socio-economic status are useful in understanding behavior insofar as they lead the researcher to investigate the particular patterns of experiences which people occupying such categories have had. Other analyses of the child-training data upon which the present study is based indicate that "only" children were judged more strongly identified with their parents, as measured by the super-ego rating, than were children of other sibling statuses (11). This finding, together with the effect of identification on aggression for boys as reported above, makes reasonable the result that the "only" boys were most aggres-

sive. One can speculate that P. S. Sears may have had some highly identified boys in her "younger" group, but the evidence is not available.

Socio-economic Status (SES)

Previous studies of doll play aggression have utilized children from widely differing social class groups. Those done in university (1, 3, 7, 13) or private preschools (4) have drawn subjects from mainly middle-class populations. P. S. Sears reported that her group (8, 12) represented "a predominantly upper-lower and lower-middle class position." Yarrow (15) had a definitely lower-class group. In the absence of any direct comparison of the aggressiveness of children from the different SES levels, it has been difficult to know how widely the results of the various studies could be generalized.

The present sample was chosen to represent a wide range of SES. Warner's Index of Status Characteristics was computed for each family on the basis of the father's income and occupational status. Maccoby and Gibbs (6) have reported in detail the rationale and computation of the score, and the SES characteristics of the present sample. For our purposes the sample was divided into two groups which may be described as "working class" and "middle class." As may be seen from Table 8, there were no differences in mean % Aggression scores attributable to SES in either doll play session.

TABLE 8

MEAN % AGGRESSION: THE INFLUENCE OF SOCIO-ECONOMIC STATUS

	BOYS			GIRLS		
	N	Session I	Session II	*N*	Session I	Session II
Middle Class	117	13.76	22.36	86	8.80	12.58
Working Class ..	82	15.28	22.48	88	8.76	13.88

SUMMARY

A test was made of the hypothesis that the frequency of five-year-old children's fantasy aggression is in part a function of the children's identification with aggressive role models. The frequency of aggression was measured in two sessions of doll play with each of 241 children, 126 boys and 115 girls. The degree of the children's identification with their parents, the severity with which the parents punished aggression, and an indication of which parent usually did the punishing were ascertained from interviews with the children's mothers.

As was predicted, boys who were highly identified, and who had the distinctive cues for male aggressiveness provided by being usually punished

by their fathers, showed the highest frequency of aggression. Identification was related to high aggression in the girls only when it was associated with severe punishment by mothers who usually did the punishing, i.e., when the girl was identified with an aggressive role model.

The present study replicated several earlier reports in finding that, in both doll play sessions, boys were more aggressive than girls, and that, for both sexes, there was an increase in the amount of aggression from the first to the second session. "Only" boys were more aggressive than boys of other ordinal positions, but there were no ordinal position differences for girls. The difference for boys was related to differences in child training practices by the parents of the "only" children. No differences in aggression were attributable to the socio-economic status of the children's families.

References

1. Bach, G. R. Young children's play fantasies. *Psychol. Monogr.*, 1945, 59, No. 2.
2. Freud, Anna. *The ego and mechanisms of defense.* London: Hogarth, 1937.
3. Hollenberg, Eleanor, & Sperry, Margaret. Some antecedents of aggression and effects of frustration in doll play. *Personality*, 1951, 1, 32-43.
4. Johnson, Elizabeth Z. Attitudes of children toward authority as projected in their doll play at two age levels. Unpublished doctor's dissertation, Harvard Univer., 1951.
5. Krall. Vita. Personality characteristics of accident repeating children. *J. abnorm. soc. Psychol.*, 1953, 48, 99-107.
6. Maccoby, Eleanor, & Gibbs, Patricia. Methods of child-rearing in two social classes. In W. E. Martin & Celia B. Stendler (Eds.) *Readings in child development.* New York: Harcourt, Brace, 1954.
7. Pintler, Margaret H., Phillips, Ruth, & Sears, R. R. Sex differences in the projective doll play of preschool children. *J. Psychol.*, 1946, 21, 73-80.
8. Sears, Pauline S. Doll play aggression in normal young children: influence of sex, age, sibling status, father's absence. *Psychol. Monogr.*, 1951, 65, No. 6.
9. Sears, Pauline S. Child-rearing factors related to playing of sex-typed roles. *Amer. Psychologist*, 1953, 8, 431 (Abstract).
10. Sears, R. R. Influence of methodological factors on doll play performance. *Child Develpm.*, 1947, 18, 190-197.
11. Sears, R. R., Maccoby, Eleanor, & Levin, H. *Patterns of child rearing*, in press.

12. Sears, R. R., Pintler, Margaret H., & Sears, Pauline S. Effect of father separation on preschool children's doll play aggression. *Child Develpm.*, 1946, 17, 219-243.

13. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, Pauline S. Some child rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-234.

14. Walker, Helen, & Lev, J. *Statistical inference.* New York: Henry Holt, 1953.

15. Yarrow, L. J. The effect of antecendent frustration on projective play. *Psychol. Monogr.*, 1948, 62, No. 6.

THE DEVELOPMENT OF PREJUDICE: A CRITIQUE OF THE *HOROWITZ FACES TEST*

EUGENE E. LEVITT

Institute for Juvenile Research, Chicago[1]

Twenty years ago Horowitz (10) published the results of a study of the development of anti-Negro prejudice in children using a "faces" test, which he devised for the purpose. The Horowitz Faces Test consists of a cardboard about 12 inches square on which are the faces of 12 boys, eight Negroes and four whites. The ages probably range from about 5 years to about 12 years. Each face is identified by a letter from A to L. The subjects are required to perform two tasks: (1) rank all the faces for preference; (2) respond to 12 sociometric questions, such as "Show me all those you want to be in your class at school," "Show me all those you like," and "Show me all those you want to sit next to in the movies."

The score on the first task, the Ranks, is the sum of the rankings of the four white boys. The possible range is from 10 (the four white boys preferred) to 42 (the four white boys least preferred.) On the second task, the Show Me, the score is the percentage of white boys of all those selected.

The original Horowitz study has never been replicated, and in fact, the present writer knows of only three researches in which the test has been used. Nonetheless, Horowitz based certain important conclusions on his work and a survey of the literature since 1935 indicates that these conclusions have been widely accepted by social scientists.[2] Briefly, the conclusions

[1] This paper was written while the author was at the Iowa Child Welfare Research Station, University of Iowa. Its preparation was partly supported by Research Grant MH-301 from the National Institute of Mental Health, U. S. Public Health Service. The author is indebted to Drs. Maurice S. Schaeffer and Boyd R. McCandless for many excellent comments and criticisms which have influenced the exposition.

[2] Horowitz' conclusions, especially the first one, are quoted in many texts in social psychology, for example, (3, 13, 16, 22, 29). They constitute the basis of a section on the development of prejudice in a recent volume on the subject (21). The first conclusion has been quoted by Hyman and Sheatsley (12) in support of criticisms of the California studies of the anti-democratic personality.

were: (1) Attitude toward the Negro is determined not by contact with the Negro himself, but by contact with the prevalent attitude toward the Negro. (2) Prejudice develops gradually, the growth curve being of the sigmoid type. (3) With increase in age from 5 to 14, an individual's attitude becomes more integrated, i.e., more consistently expressed despite differences in measuring instruments.

The purpose of the present paper is to examine critically the original Horowitz study, especially in light of other researches and findings. The critique is subdivided into four main sections concerned with (1) experimental design and statistical analysis, (2) test validity, (3) test reliability and (4) developmental theory.

Design and Analysis

The 725 subjects used by Horowitz ranged from kindergarten to grade eight, and included small samples of elementary school children from urban Tennessee and Georgia and from rural Georgia. Four hundred and seventy were from New York City schools. Sixteen had Communist parents, and 35 attended a racially-mixed school. All were boys. The experimental design is, of course, greatly simplified by having a unisexual population. It precludes the necessity for two faces tests, one for each sex, and reduces the data collection and analysis by half. Unfortunately, it also limits the conclusions, unless one accepts, *a priori,* that there will be no sex differences worth noting. In view of the studies (cf. 24, 25) which show differences in racial and religious tolerance between adult males and females, it would hardly seem prudent to make such an assumption.

The four appendices to Horowitz' article (10) contain means, correlation coefficients, and critical ratios, but they are devoid of variance data. However, the size of the means relative to the ranges appear to indicate that the distributions of both the Ranks and the Show Me scores are markedly skewed. For the 472 New York City children, the mean Ranks score is 14.9 and the range is 10–42. The Show Me has a mean of 65 and a range of 0–100. Estimated standard deviations of the two distributions for the New York sample can be derived from Horowitz' Table 1 which is concerned with regression equations and curve-fitting. The sigmas are 20.7 for the Ranks and 20.6 for the Show Me.

It is most likely that a distribution of 470 scores with a mean of 14.9, a sigma of 20.7, and a range from 10 to 42 is not normal. In all probability, the Show Me distribution is also not normal. All of the statistical techniques used by Horowitz are parametric, depending for their proper application upon underlying normal distributions. Hence they have been improperly applied. They include the various correlation coefficients, critical ratios, and comparison of curves, in short, the fundament of Horowitz' primary conclusions.

The use of inapplicable statistical techniques is a serious defect, but it should not be construed as reflecting unfavorably on the experimenter in this case. Except for chi square and rank correlation, distribution-free statistical methods were not available in 1935. Horowitz had little choice but to go ahead with the use of known techniques.

Nonetheless, we must regard Horowitz' conclusions with caution. For instance, the conclusion that attitude toward the Negro is not a function of actual contact with the Negro himself is based primarily on statistical comparisons which show that the Southern samples do not differ from the New York City sample. The actual critical ratios are not given except for the Ranks for the sixth grade. It is possible that the variability of the Southern samples is considerably less than that of the Northern sample. We do not know, but if so, a Type II error resulting from the use of a parametric comparison becomes a definite possibility. A nonparametric procedure like the median test might show a significant difference between the two distributions.

The possibility of a Type II error is heightened by the small number of cases in the Southern samples. The largest subsample is the fourth grade urban Georgia group which numbers 35. The smallest is the kindergarten Tennessee group with only 4 subjects. Of the 12 subsamples, only 3 contained more than 20 subjects. The total Tennessee sample comprises 60 subjects spread over six grades. Even with proper statistical procedures, the number of Southern subjects is too small for any definite conclusions.

Validity

If a subject were to rank the 12 faces in a random fashion, his Rank score would be 26, the midpoint of the range. If he selected a given number of faces on the Show Me in the same way, his score would be 33.33, since the white faces are one-third of the total number on the card.

A significant deviation from these chance scores indicates that some biasing factor or factors are operating. If the deviation is in the direction of a lower score than 26, or a higher percentage than 33.33, the biasing factor is race prejudice, according to Horowitz. A chance score indicates that the subject is "responding without prejudice," while a deviation above 26 on the Ranks and below 33.33 on the Show Me means that the subject is "responding with a preference for Negroes."

In attitude scaling, a high score is usually taken to mean that the subject is prejudiced. A low score means that he is unprejudiced, while a score somewhere in the middle of the distribution indicates an average amount of prejudice. There is no score which corresponds to the "preference for Negroes" although any of the low scores could be so interpreted. Most researchers in the field of race relations are not concerned with preference for Negroes as distinct from lack of prejudice. There are several good

reasons for this unconcern. One reason is that while there are many unprejudiced persons, there are probably very few white individuals who generally *prefer* Negroes to whites. This applies to children as well as adults. The present writer knows of no actual data on this point, but it appears to be a highly plausible assumption.

A peculiarity of the Faces Test is masked by the expression "a preference for Negroes." Since a low Ranks score and high Show Me score indicate prejudice against Negroes, then a high Ranks - low Show Me score actually shows a prejudice *against whites*. If one desires to hold fast to the preferences for Negroes idea, then its opposite is logically preference for whites, rather than prejudice against Negroes.

An expression like prejudice against whites (and preference for Negroes) is obviously almost meaningless when applied to a group of whites. As pointed out previously, the social scientist is disinterested in such a concept, and has no reason to try to measure it. Yet, its existence and measurement are inextricably bound up with the assumptions basic to the Faces Test.

If we simply accept the assumption that prejudice against whites is a highly improbable phenomenon in a group of white children, we can explain Horowitz' results without concluding that his group means indicate anti-Negro prejudice. Assuming that anti-white prejudice is nonexistent in the group, a Ranks score higher than 26 or a Show Me score lower than 33.33 (within certain limits of error, of course) should not occur except by chance. These two scores constitute the no-prejudice limit of the test. The two other limits are 10 for the Ranks and 100 for the Show Me. The effective ranges are, therefore, 26–10 for the Ranks, and 33.33–100 for the Show Me. The mid-points are approximately 18 and 67, and we would expect the empirical means to fall somewhere around those two scores. As it turned out, the respective means for 674 unselected Northern and Southern subjects in Horowitz' sample were about 15 and 66, which correspond fairly closely to the hypothetical averages. When viewed in this light, the Faces Test shows that little or no prejudice exists in the sample tested by Horowitz.

Following along with this analysis, what would be the proper conclusion if the empirical mean actually was 26 or 33.33? This would mean that large numbers of subjects are manifesting anti-white prejudice. Since this is *a priori* improbable, the test must be at fault. In the absence of any extenuating evidence, we would have to conclude that the test is invalid. Horowitz' assumptions are completely twisted; the results which he calls indicative of prejudice are actually those to be expected purely by chance, while a realization of what he feels is the chance expectation would nullify his test.

Suppose we retrace several paragraphs and assume for the moment that Horowitz' assumptions are tenable. A group which yielded the mean scores to be expected by chance is unprejudiced. It must also be expected

that the scores of this group will be distributed in some way around the respective means. The theoretical chance distributions are easily determined. For the Ranks, for instance, there are 495 possible scores[3] falling symmetrically about the mean of 26. The sigma of this distribution is 5.89. If we assume that the group has responded to the test in a purely random fashion, this is approximately the distribution to be expected. Statistically speaking, this distribution includes the scores of individuals who are probably prejudiced, viz., the 12 scores falling more than two sigmas below the mean (less than 14.22). That is, the fact that the *group* is unprejudiced does not in this case preclude the possibility of *prejudiced individuals* within the group. Since the group is ranking the faces randomly, we cannot be sure just which individuals are the prejudiced ones. Assuming adequate test reliability, a retest should yield the same mean score, but individual scores might change considerably. However, any subject whose personal chance distribution includes a score two sigmas below the mean ought to be suspect. At any rate, it is reasonably sound statistically to say that an individual who scores more than two sigmas below the mean is prejudiced. Indeed, it is a necessary statement if any individual prediction is ever going to be attempted.

Now suppose that we obtain a distribution with a mean of 18 and a sigma of 1.1. The mean differs significantly from 26, but the distribution does not include a single score lower than 14.22. Conclusion: the group is prejudiced, but it does not have a single prejudiced individual in it! It is obvious that we must insist that the variances of the empirical and the theoretical distributions be roughly the same if this paradox is to be avoided. But this is not likely if the mean is 18 or thereabouts. The lower limit of the range is 10 which would leave base-line room for less than 1.5 sigma below a mean of 18. If the sigma is 5.89, the distribution is going to quite skewed. A parametric comparison of the obtained and theoretical means is hence inappropriate. A satisfactory nonparametric comparison like the median test might not reveal a significant difference since it is probable that the medians of the obtained and theoretical distributions are closer together than the means.

The use of deviation from chance as a criterion of test validity is thus questionable. Even if this method stood up under logical analysis, the validity of the Faces Test as a measure of prejudice would be far from demonstrated. The test results might differ from chance due to the functioning of a biasing factor, but it would still remain to be shown that this factor was what we call race prejudice. The Faces Test directly measures *preference* for whites over Negroes. This alone is not commonly regarded as bigotry. Logically it is quite a leap from preference for, to prejudice against. Preference for one kind of stimulus does not necessarily imply hostility toward a

[3] The theoretical distribution of Ranks scores is obtained by means of a formula for all possible combinations of 4 ranks out of 12 ranks. This formula is 12! / (4!) (8!).

second kind. As Allport (1) points out, the concept of hostility is a basic one in a definition of race prejudice. Horowitz made no real attempt to demonstrate that responses to the Faces Test were motivated by hostility toward the Negro. He recorded spontaneous comments of three children, one Northern and two Southern, to show that they were "recognizing the racial aspect of the tests," (the Northern child is a 4-year old girl) and 7 unclassified comments indicating that the children were "responding realistically." Even if there were a sufficient number of such remarks by representative members of the experimental population, their merit as evidence for validity would be doubtful.

In 1950, Mussen (18) published the results of a study using the Faces Test as a measure of anti-Negro feeling. He reported (17) that there was a correlation of .52 between Faces scores and choice of cabin-mates by white boys in a racially-mixed summer camp in the Northeast. Mussen derived a single Faces score by arbitrarily weighting the Show Me score five times and combining it with the Ranks score by means of standard scores, a procedure which practically eliminates the Ranks from consideration. The relationship between the Faces Test and choice of cabin-mates indicates that the former is valid as a sociometric indicator for Northern boys. The general results (18) are in accord with theoretical considerations, and we may accept Mussen's findings as indicating a degree of validity for the Faces Test. This is, however, the only such instance in the literature.

The Mussen study is one of the three published correlational researches in which the Faces Test was utilized as an index of prejudice. The others (14, 15) report negative relationships between the Faces Test and other measures. Levitt and McCandless (14) reported that both the Ranks and the Show Me were essentially unrelated to two conventional attitude scales for the measurement of anti-Negro feeling in children. McCandless and Holloway (15), using the Mussen combined score, found no correlation between the Faces Test and six measures of intolerance of ambiguity. The samples in both studies consisted of boys and girls in intermediate elementary school grades in two medium-sized Midwestern towns. It is possible, of course, that the Faces Test might be related to attitude or intolerance of ambiguity measures in populations in other sections of the country. In any event, the absence of relationships (which would be predicted to exist on theoretical grounds) in at least one kind of population limits the generality of conclusions stemming from the Faces Test.

Both Mussen (17) and Horowitz (11) are of the opinion that the Faces Test should produce different results than an ordinary attitude scale because of differences in the test structure, and because each test is likely to be tapping a different subarea of a nonunitary trait. This explanation could account for a low positive correlation between the Faces Test and the attitude scale, but it falls short of explaining an absence of relationship. Different subareas of a characteristic ought to be somewhat related, other-

wise there is no empirical basis for regarding them as part of the same characteristic.

Certain other findings indicate that the preference for whites expressed on the Faces Test need not spring from hostility toward Negroes. It is reasonable to expect that children, with their burden of insecurity in a complex world, will prefer the familiar so long as it is unthreatening. None of the faces of Horowitz' test are personally known to the subject, but white skin is more usual in his private milieu, and hence more familiar in the general sense. That characteristic which most resembles himself becomes an important basis for preference where other bases are lacking.

Several studies support this reasoning. Springer (27) has shown that preschool children of three different racial groups manifested a modal preference for faces of their own race. There were no differences in the degree to which each group preferred its own faces. In a study by Clark and Clark (4) it was found that identification with a Negro or a white face by Negro preschool children was a function of the lightness or darkness of the responding child. Fifty-six per cent of the very dark children, but only 36 per cent of the light-skinned Negro children, identified with the Negro face. This seems to be a clear-cut instance of preference for the self-resembling face, regardless of other considerations.

The Clark and Clark experiment also gives another indication that preference on a faces-type test is probably unrelated to hostility. As the Clarks note, the responses of the light-skinned children suggest that the identifications were made "not on the basis of socially defined racial group differences, but on the basis of physical characteristics. . . ." (4, p. 160). The data from the Springer study (27) are in accord with the Clarks' conclusion. When asked for reasons for preference, Springer's children gave physical characteristics in 74 per cent of the instances. "Looks like me" made up more than 9 per cent of the reasons, while nationality was referred to in only 3 per cent of the reasons. Helgerson (9) found that when subjects and pictures of both sexes were used, the sex of the pictured face is more important than race in determining choice of prospective playmates.

It is true that all these studies used preschool children as subjects while Horowitz' subjects were grade school pupils. Furthermore, it is possible that an overtly expressed reason may be at variance with underlying motivation. Nonetheless, these findings are highly suggestive. They indicate that it is quite possible that children's preferences for faces are based on similarity to themselves, or upon more superficial characteristics than race *per se*. If this is true, it seems unlikely that the preference for white faces manifested by white subjects responding to the Horowitz Faces Test is motivated by hostility toward the Negro faces.

Some additional light on this point is cast by an analysis of the valences of the individual faces in Horowitz' test. By designed controlling of albedo,

the Faces Test is arranged so that there are four light-skinned Negro faces and four dark-skinned Negro faces. In the alternate form of the test, the albedos are reversed, i.e., the light-skinned faces are dark, and the dark-skinned faces are light. By comparing group preferences for individual faces on the two forms, Horowitz concluded that preference for the lighter Negro faces exists only at the kindergarten and first grade levels.

However, an analysis of Faces Test data from a sample of 76 boys tested by Levitt and McCandless (14) fails to confirm Horowitz' finding. The approach was an analysis of variance into individual degrees of freedom using the method described by Snedecor (26). Using this procedure it is possible to determine whether or not there are significant differences between the light-skinned and the dark-skinned Negro faces each taken as a group. This analysis was done for both the Ranks and the Show Me.[4] For the Show Me, the data were the absolute number of times the face was selected per subject.

The resulting *F*-ratios are 7.25 for the Ranks, and 9.60 for the Show Me, both significant beyond the .01 level. In each case, the light-skinned group is the preferred one. The mean Ranks scores are 6.69 for the light-skinned group and 7.42 for the dark. The respective Show Me scores, i.e., mean number of times chosen per subject, are 5.57 and 4.81. These results appear to indicate rather definitely that in this sample of fifth grade subjects, albedo of the Negro faces is a factor determining preference.

An interesting side light is revealed by the analysis into individual degrees of freedom. In addition to the differences already noted, the light-skinned Negro faces, the dark-skinned Negro faces, and the white faces all differ significantly *among themselves* for both the Ranks and the Show Me. That is, some of the faces in each group are preferred to others in the same group. If we rank the mean scores per face on each test, a rank order correlation of .94 is obtained, indicating that the relative valence of each face is consistent from test to test.

The unequivalence of faces within groups raises the question of possible bias due to the stimuli themselves. An expressed preference of some kind for Negro face B (mean Ranks score, 4.58, mean Show Me score, 7.86) may mean something quite different from a preference for Negro face G (mean Ranks score, 8.54, mean Show Me score, 4.10). The same consideration applies to white face L (mean Ranks score, 4.72, mean Show Me score, 7.96) and white face F (mean Ranks score, 6.57, mean Show Me score, 6.19). If we accept the basic hypothesis that prejudice can be measured by preference for faces, then we would expect that an individual who selected the single Negro face, B, would be more prejudiced than one

[4] Technically speaking, Snedecor's procedure is not appropriate for ranked data. The writer knows of no method for breaking down the variance of ranked data into individual degrees of freedom. However, we may assume that Snedecor's method gives an approximate result.

who chose the single Negro face, G. Rejecting face L would indicate less prejudice than rejecting face F. The usual method of scoring the Faces Test does not consider such differences, but simply treats all of the white faces, and all the Negro faces as equivalent. It might be added parenthetically that if the Faces Test is internally consistent, we would expect a high degree of consistency of response to individual faces. This may be true of the Show Me, but a rough inspection of the data indicates that it will hardly hold for the Ranks. For instance, face L, which has lowest (most preferred) mean Ranks score of the whites, was ranked above all the other white faces by only 18 of the 76 subjects.

The Show Me score is simply the percentage of white faces of those selected, irrespective of the absolute number of white faces chosen. The highest possible Show Me score, 100, could be obtained by choosing a single white face and no Negro faces, or by choosing 48 white faces (the maximum possible) and no Negro faces. It seems indisputable that a subject following the latter procedure is the more "prejudiced" of the two, though the test will not distinguish between them. We ought to find, then, that the tendency to select white faces is inversely proportional to the tendency to select Negro faces. The more white faces a subject selects, the more prejudiced he is, hence he should select fewer Negro faces.

This adds up to a negative correlation of some degree between the number of white and the number of Negro faces chosen. Actually, Levitt and McCandless (14) found just the opposite. For the sample of 76 Midwestern fifth grade boys, there was a *positive* correlation of .49 between the number of white faces and the number of Negro faces chosen on the Show Me. For a comparable sample of 78 girls, the correlation was .58.

This seems to indicate that responses to the Show Me are determined to some extent by a factor which we might call "accepting strange faces." This factor may, or may not, confound the Show Me results, but at any rate, it is something different from "accepting Negro faces," which was evidently assumed to be the key factor underlying responses to the Show Me.

A common method of establishing the validity of a new test involves the use of contrasting criterion groups. Suppose it is known or hypothesized on good theoretical grounds that the members of Group A have more of variable X than the members of group B. If the test shows, in fact, that this is the case, we conclude that it has some validity as a measure of variable X. Contrariwise, if Group A does no better, or worse, than Group B, we are led to suspect that the test is invalid as a measure of variable X. Of course, the test constructor can always fall back on a stubborn operationism, and insist that not only is his test valid, but that his results demonstrate that the supposed superiority of Group A is an "oft-repeated cliché" with no foundation in fact. The tenability of his claim will rest largely on the soundness of the evidence supporting Group A's superiority.

Not many social psychologists or sociologists would deny that there is a greater prejudice against the Negro in the South than in the North. As Horowitz himself once put it: "Most attempts to analyze attitude toward Negroes as a single, generalized unit suggest that Southern samples are least friendly, most 'prejudiced,' most reactionary, least tolerant, that the Northern samples are the reverse" (11, pp. 206-207). It is quite possible that Horowitz originally used the contrasting Northern and Southern groups in his study in an effort to show that the Faces Test could discriminate between them, viz., that the Southern group would show more prejudice. When it turned out that the groups did not differ (we shall assume for the moment that proper statistical procedures would lead to the same conclusions as improper ones) Horowitz faced alternate conclusions: (1) that the Faces Test is invalid as a measure of anti-Negro feeling; (2) that Southern white children are really not more prejudiced against Negroes than Northern white children. Obviously, he chose the latter, and surprisingly, many social scientists went along with him.

Horowitz used other contrasting groups in an effort to demonstrate test validity, and to support his conclusions. His small sample of children of Communists have a considerably higher Ranks mean and a lower Show Me mean than the other Northern white boys. The purpose of the inclusion of this group is obvious: Communists have a more favorable attitude toward minority groups than other people. Since that is what the test shows, in fact, then the test is valid. But the contrasting groups kind of validity, as pointed out previously, depends upon the soundness of evidence supporting the hypothetical contrast. It will be agreed, no doubt, that Communists *manifest* a greater acceptance of Negroes, but it is far from certain that this acceptance stems from genuine tolerance. As Shils (23) has pointed out in his penetrating analysis of the California studies of the anti-democratic personality, the leftist, like the rightist, is characterized by a rigid adherence to a particular set of stereotyped beliefs. The underlying personality structures may be similar, though the clichés differ. The Communist who conforms to his code *must ostensibly* accept minority group members, even though he seethes with suppressed hostility toward them. This prescribed credo of tolerance makes the Communist sample a poor one for contrasting groups validity.

Thirty-five sixth grade subjects from a mixed school were included in Horowitz' testing. Eleven of these were white children. Their mean Faces Test scores do not appear to differ from those of the children in the remainder of the New York City sample.

Horowitz therefore regarded this finding as further evidence that anti-Negro feeling is not determined by contact. The contrast is enervated by the small *N* and inappropriate statistical analysis, but more importantly, it conflicts with expectation based on a number of other researches. Negro children formed about 70 per cent of Horowitz' class. Criswell (5), using

a sociometric technique, demonstrated that a few Negro children in a predominantly white class may be regarded as pets, while a large Negro representation is seen as a distinct threat. We would expect, therefore, that Horowitz would have found more anti-Negro feeling in this class than in others if his test was a valid index. One can argue that it is Criswell against Horwitz, but other data support the former's findings. For example, in the small-town Midwestern population sampled by Radke and Sutherland (19), prejudice did not appear until about age 11 or 12. The Negro population in the community was tiny. In the samples from large, Northeastern urban areas (where the proportion of Negroes is much higher) tested by Trager and Yarrow (28) and by Goodman (6), there was an appreciable smattering of anti-Negro feeling at the preschool level. Forty per cent of a New England sample (2) reported a first experience of anti-Negro feeling by age 11, while only 20 per cent made the same report in South Dakota (20).

This is another instance of the use of contrasting groups, and there are the same possibilities. Such evidence as there is indicates that the white children in the mixed class ought to show more hostility. According to the Faces Test, they do not. Either the assumption is incorrect, or the test is invalid. It appears that the scales tip in favor of the latter conclusion.

Reliability

Mussen (17) reports an estimated split-half reliability of .94 for the Show Me. No coefficient is given for the Ranks, for some unknown reason. Horowitz does not report any statistical estimates, but gives the correlations by grade between two forms of both tests, and test-retest correlations for four grades after a six-months' period. The average correlation between the two forms is .68 for the Ranks, and .79 for the Show Me. The respective average test-retest coefficients are .59 and .50.

Since little data on the second form of the tests are offered, it is not possible to evaluate the correlation between forms. The test-retest correlations are, of course, extremely low, but this may be a function of the unusually large interval between testings. Especially with children, such an hiatus is apt to attenuate the coefficient.

The reliability found by Mussen appears satisfactorily high. However, the high split-half correlation may result in part from the fact that a sizeable proportion of subjects have selected all or nearly all 144 faces in response to the 12 Show Me items. In the Levitt-McCandless experiments, almost 10 per cent of the boys chose all 144 faces, while an additional 10 per cent chose 131 faces or more. Naturally, the split-half reliability of this 20 per cent of the results is going to be very high, probably .90 or better, and will markedly influence the reliability in the whole group. It is also possible that the percentage of Mussen's sample choosing all 144 faces was greater than in the Levitt-McCandless work.

As a possible check on the figures given by Horowitz and Mussen, split-half reliabilities were computed from the scores of the boys in the Levitt-McCandless study. The Ranks coefficient was obtained by correlating the sum of the rankings for the white faces C and F, with the sum of the rankings for white faces I and L. The Guttman coefficient (7) for these data is .48. The reliability for the Show Me is .89, an estimate obtained by correlating the percentage of white faces chosen on the first six questions with that on the last six. The reliabilities obtained from the Levitt-McCandless data are approximately the same as the Show Me coefficient found by Mussen and the test-retest reliability for the Ranks reported by Horowitz. We may conclude, at least tentatively, that the reliability for the Ranks is around .50, for the Show Me, about .90.

There is a connection between the Ranks and Question 12 of the Show Me—"Show me all those you like"—which permits an interesting and unusual investigatiton of reliability. Since the Ranks test is a ranking of the faces for liking, it ought to "scale," so to speak, with Question 12. That is, the faces which are selected in response to Question 12, should be found consecutively ranked from the top on down in the Ranks, with no gaps. A subject who chooses faces B, D, F, G, and K on Question 12 should also place these five faces in the first five rankings, the order being immaterial. It would be inconsistent, to say the least, if a subject ranked any face which he did not like ahead of those which he did like. In this sense, the Ranks and Question 12 are equivalent stimuli.

Analysis of test results show that inconsistency occurs more often than not. Of the 76 fifth grade boys, 42, or 56 per cent, had rankings inconsistent with Question 12. Of the 34 whose rankings were consistent, 25 had selected all of the faces in response to Question 12, (that is, it was impossible for their rankings to be inconsistent). Of the remaining nine, three had selected 11 of the 12 faces, which does not leave much room for inconsistency.

In other words, the stated preferences show a marked unreliability, even though the interval between presentations of the equivalent forms is only a few minutes. This raises a question of the significance of the choice of a face by a grade school child. Can the child actually be expected to express a meaningful preference among individuals merely by observing a face? It might be argued that inconsistency *per se* does not make for unreliability of the test as a measure of anti-Negro feeling. This argument is likely to be specious since in all but eight of the 42 inconsistent rankings, the deviation in rankings from the selections of Question 12 involved changes in position of a white face or faces relative to the Negro faces. That is, in each of these 34 instances, the Ranks score would change if it were made consistent with Question 12.

The lack of equivalence of the Ranks and Question 12 together with the relatively low Ranks reliability suggest that the Faces Test is not a particularly reliable instrument.

Developmental Theory

Horowitz developed three "theoretical growth curves of attitude" from his data.[5] These curves are the basis of his conclusion that prejudice develops gradually, following an S-shaped pattern. They are reproduced in Figure 1.[6] In selecting the curve for the Show Me as the representative

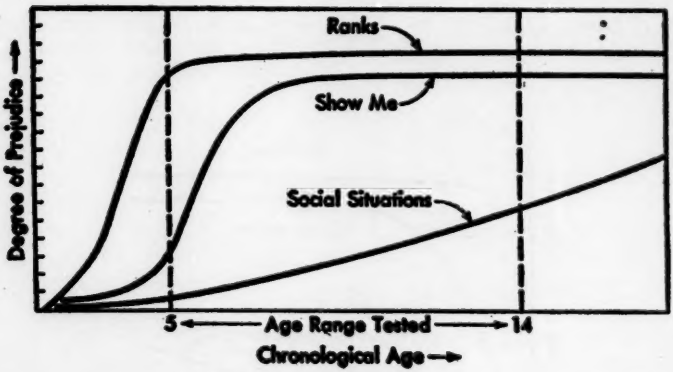

Figure 1—Theoretical growth curves of attitude toward the Negro according to Horowitz.

pattern, Horowitz says that "the three curves derived in this study may merely represent segments from different parts of three S-shaped curves" (10, p. 25). This speculation is, of course, a possibility, but the curves in Figure 1 are somewhat misleading. The actual data begin at the dotted line marking off age 5. The parts of the curves to the left of the dotted line are extrapolations, not based on obtained data. Figure 2 shows the curves stripped of extrapolations. It can be seen that there is nothing sigmoid about any of the curves, and that, of the three, only the Social Situations has a gradually rising slope. (Even the little slope at the beginning of the Ranks curve is unwarranted according to Horowitz' analysis of the data.)

[5] The third curve represents the Social Situations Test, a pictorial-type instrument similar in design and intent to the Faces Test. It was also used in Horowitz' original study (10). It is more or less independent of the Faces Test having an average correlation of .16 with the Ranks and .23 with the Show Me. Its test-retest reliability over the six-months interval was only .16.

[6] Figure 1 is a reproduction of a slightly modified version of Horowitz' original figure. It is taken from the recent text by Hartley and Hartley (8, p. 704).

The conclusions concerning development which appear warranted by the curves are: (1) Attitude as measured by the Ranks shows no real development. The attitude is completely developed at age 5, and does not change in succeeding years. (2) Attitude as measured by the Show Me develops rapidly from age 5 on, and is completely developed by age 8 or 9. After that, it does not change. (3) Attitude as measured by the Social Situations develops gradually from age 5 on, is still developing at age 14, and may conceivably develop further after that age. There is no statistical justification for choosing any one of the three as the dominant pattern, nor does there seem to be any logical way in which the three curves can be fused.

FIGURE 2—Horowitz' theoretical growth curves with extrapolations removed.

While it is true there is no statistical reason for looking upon one of the three curves as the curve, the Social Situations test has the most in its favor rationally. It is the only one which shows a continuing development over the grade school period. The Show Me curve could conceivably begin to ascend again after age 14, but it does not seem logical that there should be an attitudinal vacuum for the five preceding years.

In view of the defects of the test and the faulty statistical procedures, one cannot make any statement concerning the development of prejudice based on the Horowitz Faces Test. To be on fairly safe grounds, a conclusion concerning development should be based on a reasonably valid and reliable test, a cross-sectional sample of both sexes from various parts of the

country, and proper statistical methodology. Such bases are not yet available, so far as the present writer knows.

Summary

The Horowitz Faces Test and the original work with it (10), which have been the primary basis for significant conclusions concerning the development of anti-Negro prejudice in children, were critically examined. Defects in research design and analysis, principally the use of inappropriate parametric statistics, were found. Analysis indicates that validity of the Faces Test is unsatisfactory for a number of reasons: (1) The use of deviation from chance as evidence of prejudice is logically untenable. (2) The test measures preference for whites or Negroes, but it has not been shown that these preferences spring from hostility. Contrary evidence is presented by the writer. (3) Two of three correlational studies report no relationships between the Faces Test and tests to which a valid prejudice measure should relate on theoretical grounds. (4) The various faces of the test are not equivalent stimuli, though the scoring of the test assumes that they are. (5) An unexpected positive correlation between number of Negro faces and number of white faces chosen has been found. (6) Evidence from contrasting groups used by Horowitz to support his conclusions is questionable.

Reliability of the Ranks was found to be unsatisfactory, though reliability of the Show Me was adequate. It was shown that subjects tended to respond differentially to the Ranks and to a very similar stimulus in the Show Me. The question of the meaning of a preference for a face by a child is raised.

Further analysis indicates that Horowitz' conclusions concerning growth of attitude are unsupported by his own data. None of the obtained curves is actually sigmoid as he claimed.

It is concluded that in view of the various shortcomings of the Faces Test and Horowitz' experiment, no definite statements concerning the development of prejudice are warranted.

References

1. Allport, G. W. *The nature of prejudice.* Cambridge, Mass.: Addison-Wesley, 1954.
2. Allport, G. W., & Kramer, B. M. Some roots of prejudice. *J. Psychol.*, 1946, 22, 9-39.
3. Bird, C. *Social psychology.* New York: Appleton-Century, 1940.
4. Clark, K. B., & Clark, Mamie K. Skin color as a factor in racial identification of Negro preschool children. *J. soc. Psychol.*, 1940, 11, 159-169.

5. Criswell, Joan H. A sociometric study of race cleavage in the classroom. *Arch. Psychol.*, 1939, No. 235.

6. Goodman, Mary E. *Race awareness in young children.* Cambridge, Mass.: Addison-Wesley, 1952.

7. Guttman, L. A basis for analyzing test-retest reliability. *Psychometrika,* 1945, 10, 255-282.

8. Hartley, E. L., & Hartley, Ruth E. *Fundamentals of social psychology.* New York: Knopf, 1952.

9. Helgerson, Evelyn. The relative significance of race, sex, and facial expression in the choice of playmates by the preschool child. *J. Negro Educ.*, 1943, 12, 617-622.

10. Horowitz, E. L. The development of attitude toward the Negro. *Arch. Psychol.*, 1936, No. 194.

11. Horowitz, E. L. Race attitudes. In O. Klineberg (Ed.), *Characteristics of the American Negro.* New York: Harper, 1944.

12. Hyman, H. H., & Sheatsley, P. B. The authoritarian personality—a methodological critique. In R. Christie & Marie Jahoda (Eds.), *Studies in the scope and method of the authoritarian personality.* Glencoe, Ill.: Free Press, 1954. Pp. 50-122.

13. Katz, D., & Schanck, R. L. *Social psychology.* New York: Wiley, 1938.

14. Levitt, E. E., & McCandless, B. R. Some relationships among tests pertaining to the authoritarian personality syndrome in children. *Amer. Psychologist,* 1954, 9, 417. (Abstract)

15. McCandless, B. R., & Holloway, H. D. Race prejudice and intolerance of ambiguity in children. *J. abnorm. soc. Psychol.*, 1955, 51, 692-693.

16. Murphy, G., & Murphy, Lois B. *Experimental social psychology.* (Rev. Ed.) New York: Harper, 1937.

17. Mussen, P. H. The reliability and validity of the Horowitz Faces Test. *J. abnorm. soc. Psychol.*, 1950, 45, 504-506.

18. Mussen, P. H. Some personality and social factors related to changes in children's attitudes toward Negroes. *J. abnorm. soc. Psychol.*, 1950, 45, 423-441.

19. Radke, Marian Y., & Sutherland, Jean. Children's concepts and attitudes about minority and majority American groups. *J. educ. Psychol.*, 1949, 40, 448-468.

20. Rosenblith, Judith A. A replication of "some roots of prejudice." *J. abnorm. soc. Psychol.*, 1949, 44, 470-489.

21. Saenger, G. *Social psychology of prejudice.* New York: Harper, 1953.

22. Sherif, H. *An outline of social psychology.* New York: Harper, 1948.

23. Shils, E. A. Authoritarianism: "Right" and "Left." In R. Christie & Marie Jahoda (Eds.), *Studies in the scope and method of the authoritarian personality.* Glencoe, Ill.: Free Press, 1954. Pp. 24-49.

24. Sims, V. M., & Patrick, J. R. Attitude toward the Negro of Northern and Southern college students. *J. soc. Psychol.,* 1936, 7, 192-304.

25. Smith, F. T. An experiment in modifying attitudes toward the Negro. *Teach. Coll. Contr. Educ.,* 1943, No. 887.

26. Snedecor, G. W. *Statistical methods.* (4th Ed.) Ames, Iowa: State College Press, 1953.

27. Springer, Doris V. Awareness of racial differences among preschool children in Hawaii. *Genet. Psychol. Monogr.,* 1950, 41, 215-270.

28. Trager, Helen G., & Radke, Marian Y. *They learn what they live.* New York: Harper, 1952.

29. Vaughn, W. I. *Social psychology.* New York: Odyssey Press, 1948.

THE EFFECT OF CONCEPT KNOWLEDGE ON DISCRIMINATION LEARNING[1]

WINIFRED O. SHEPARD and MAURICE SCHAEFFER[2]
Iowa Child Welfare Research Station

The purpose of the present paper is to describe a group technique for the experimental study of learning processes in children and to report on a study which employed the technique. The study was concerned with the effect of knowledge of a concept name (triangle) on the speed of learning a discrimination problem in which the correct response was always an instance of that concept. The prediction was made that *Ss* who knew the concept name would perform better than *Ss* who did not.

SUBJECTS

One hundred and seventy-eight sixth grade pupils from six classes in the Cedar Rapids, Iowa, public school system were subjects in the experiment. The classes contained 22, 31, 28, 35, 31 and 31 pupils respectively.

MATERIALS

Twenty stimuli, five instances of each of four classes of geometric figures, were used. There were five triangles, five quadrilaterals, five pentagons and five nonsense figures. These 20 different figures were arranged, as described below, in mimeographed booklets. Each booklet contained 25 pages which constituted the 25 trials of the learning task.

Each page was divided into four 2 × 2 inch compartments labeled A, B, C and D respectively. One triangle, one quadrilateral, one pentagon and one nonsense figure were printed on each page. Thus, each of the 20 geometric figures appeared in the booklet five times. The order of appear-

[1] This investigation was supported by a research grant (MH-658C) from the National Institutes of Health, U. S. Public Health Service.

[2] The authors wish to express their thanks to Mr. H. C. Reid, Director of Curriculum and Instruction of the Cedar Rapids Public Schools, and to the teachers and pupils who cooperated in this investigation.

ance of stimuli was carefully controlled so that there was no tendency for any figure or any class of figures to be systematically associated with a particular position on the pages.

In addition, a concept name test was prepared. This was a mimeographed sheet containing 18 different geometric figures, five of which were triangles.

Procedure

In each class, *E* passed out the booklets and then gave, in essence, the following instructions. "On each page of the booklet you have been given, you will find four figures. Only one of the figures on each page is the correct one. At first you will just be guessing, but after a while you will be able to tell each time which is the correct one. When I say 'ready,' turn to page one and mark the figure you think is correct. When you have done so, raise your hand. When everyone has raised his hand, I will call out the letter of the compartment that contained the correct figure. Then you will turn to page two and do the same thing." This procedure was gone through for 25 trials. The correct choice on each trial was the triangle. As a result of preliminary experimentation with this population, it was found possible to prevent *Ss* marking the pages after the correct answer had been called out merely by having *E* walk around the room randomly checking booklets on each trial. Larger classes might require additional proctoring.

Following the learning task, the booklets were collected and the concept name test was passed out. The instructions were simply to write the name of each figure in the space underneath it.

Results

The first step in the analysis was to separate the *Ss* into those who knew and did not know the concept name "triangle." Arbitrarily, *Ss* who wrote the word "triangle" under each of the five triangles that appeared on the tests were considered to know the concept. *Ss* who did not use the word at all or who used it for less than all of the instances were considered not to know the concept. According to this criterion, 58 *Ss* knew the concept and 120 did not.

Learning was defined as the occurrence of seven successive correct responses. The probability of seven or more successive correct responses in 25 trials in a situation where there are four possible choices on each trial is .0009 (3). Fifty-four or 93 per cent of those who knew it succeeded in learning the discrimination task. Eighty-seven or 73 per cent of those who did not know it learned. The probability of occurrence of this difference in proportion is less than .001.

In any situation in which *Ss* are arbitrarily broken down into two or more groups on the basis of one variable, it is wise to check on the possi-

bility that another variable may be systematically associated with the breakdown and be itself responsible for the observed difference in criterion scores. Accordingly *Ss'* scores on the Otis Self Administering Test of Mental Ability were studied. A two factor analysis of variance (6) of these IQ scores was carried out in which knowledge of the concept constituted one factor and classes the other. The results, corrected for disproportionality (10), showed that neither the concept knowledge factor nor the interaction of that factor with classes approached significance, although there was a significant difference ($.025 > p > .010$) in IQ among the classes. In view of these results, the hypothesis that knowledge of the concept name per se was responsible for the observed difference in criterion scores remains tenable.

One further analysis was carried out to see if still another response measure would reveal the superiority of those who knew the concept over those who did not and to check on the possibility of an interaction between concept knowledge and IQ. This was a 2×4 treatments by levels analysis of variance (6) in which knowledge of the concept was the treatment variable and IQ the levels variable, with total number of errors on the learning task as the criterion measure. Four IQ levels were used: 90 and below, 91–105, 106–115, and 116 or over. Due to extreme disproportionality, it was decided to eliminate randomly 16 *Ss*. Hence this analysis, which appears in Table 1, is based on 162 *Ss*. From the table, it can be seen that the treatment variable was highly significant, while neither IQ nor the interaction of IQ with concept knowledge was significant.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF ERROR SCORES

Source	*df*	*Mean Square*	*F*
Concept Knowledge	1	226.67	8.64*
IQ Level	3	30.93	1.18
Concept × IQ	3	18.56	
Within cells	154	26.23	

* $p < .005$.

DISCUSSION

These results can be explained in terms of the notion of mediated generalization based on response-produced stimulation. This is a principle that has been used to account for the fact that *Ss* who have learned one common response to a group of stimuli subsequently have increased tendencies to generalize other responses among them. The theoretical *S–R* paradigm for this phenomenon appears in Figure 1.

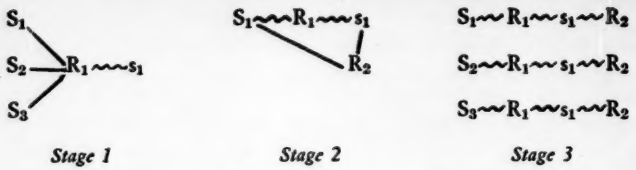

FIGURE 1—S–R paradigm for mediated generalization

S_1, S_2 and S_3 represent three distinctive stimuli. In Stage 1, *S*s are taught to make response R_1 to each of them. The occurrence of R_1 produces proprioceptive stimulation, s_1, which is assumed to have all the functional properties of any external stimulus. In Stage 2, *S*s are taught a second response, R_2 to S_1. By virtue of the training in Stage 1, S_1 is already known to elicit R_1 with its attendant s_1. During this stage, a connection is established between s_1 and R_2. Now when S_2 and S_3 are presented during Stage 3, R_2 is observed to occur although it had never been specifically conditioned to these stimuli. It occurs because both S_2 and S_3 elicit R_1 and hence s_1, and s_1 elicits R_2. It can also be seen that the Stage 3 stimulus complexes, S_1-s_1, S_2-s_1 and S_3-s_1 are more similar to each other than are S_1, S_2 and S_3 alone which appear in Stage 1. Hence generalization tendencies are higher in Stage 3.

In the paradigm, R_1 is referred to as the mediating response. Theoretically, any common response learned to a set of stimuli could serve as a mediating response. Birge (1) and Jeffrey (5) have, in fact, shown that both verbal and motor responses can be mediators.

In the present study the verbal response "triangle" served as a mediating response. Presumably, prior to the experiment some *S*s had learned to associate this response with every instance of a closed three-sided figure. During the learning task this verbal response produced proprioceptive stimulation to which the marking response became conditioned. After the first correct marking response was made an association was begun between the stimulation produced by the word "triangle" and the response of marking the figure which elicited that name. This association became rapidly stronger with each succeeding correct response. As soon as these *S*s turned a page, one stimulus on that page elicited the word "triangle" and hence the marking response. The problem for these *S*s had been learned.

On the other hand, *S*s who did not use the word "triangle" at all, or who used it for only some types of triangle, were faced with a different situation. They could learn the task in one of two ways. One way would be to simply learn the marking response separately to five different stimuli—the five triangles which appeared throughout the booklet. [Gellerman (2) showed that both monkeys and very young preverbal children could learn the con-

cept triangle in this way.] Or, these *Ss* could formulate for themselves, during the learning task, some common response such as "three-sided" which could serve to mediate correctly the generalization of the marking response. Either of these methods would take longer than the immediate application of an already available mediating response. Hence, in comparing the performance of *Ss* who did and did not know the verbal response "triangle," a larger proportion of the latter did not learn at all in the time allowed and, in general, they made more errors.

It should be noted that these findings are consistent with those of other studies, Heidbreder (4), Long and Welch (7, 8), Roberts (9), Snygg (11) and Welch and Long (12) in which it was found that children who can verbalize the basis of learning show superior performance to those who cannot, even in cases where nonverbalizers eventually do learn.

A final word remains to be said about the particular technique of gathering data employed. It is, of course, possible with this group technique to gather in a few hours the same amount of data that it takes months to gather when working with *Ss*, particularly children, on an individual basis. It is furthermore felt that the technique is amenable to the study of many discrimination learning variables. For instance, the similarity of the stimuli to each other, their number and complexity could be systematically varied. Also, motivational variables could be manipulated through the use of printed instructions for *Ss* old enough to read.

Summary

A group technique was used in studying the effect of concept name knowledge on learning in sixth grade children. The task involved choosing the triangle on every trial from among four geometric figures. There were five different instances of triangles throughout the task. Children who demonstrated on a concept name test that they could consistently and correctly use the word "triangle" were found to perform better on the learning task than children who could not. The results were discussed in terms of the mediated generalization hypothesis.

References

1. Birge, Janet. The role of verbal responses in transfer. Unpublished doctor's dissertation, Yale Univer., 1941.
2. Gellerman, L. W. Form discrimination in chimpanzees and two year old children. *J. genet. Psychol.*, 1931, 39, 50-72.
3. Grant, David A. Additional tables of the probability of "runs" of correct responses in learning and problem solving. *Psychol. Bull.*, 1947, 44, 276-279.

4. HEIDBREDER, E. Reasons used in problem solving. *J. exp. Psychol.*, 1927, 10, 397-414.

5. JEFFREY, W. E. The effects of verbal and nonverbal responses in mediating an instrumental act. *J. exp. Psychol.*, 1953, 45, 327-333.

6. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.

7. LONG, L., & WELCH, L. Reasoning ability in young children. *J. Psychol.*, 1941, 12, 21-44.

8. LONG, L., & WELCH, L. Influence of levels of abstraction on reasoning ability. *J. Psychol.*, 1942, 13, 41-59.

9. ROBERTS, K. E. Ability of preschool children to solve problems in which a principle of relationship is kept constant. *J. genet. Psychol.*, 1932, 40, 118-133.

10. SNEDECOR, G. W. *Statistical methods*. Ames: Iowa State College Press, 1946.

11. SNYGG, D. Relative difficulty of mechanically equivalent tasks. I. Human learning. *J. genet. Psychol.*, 1935, 47, 299-320.

12. WELCH, L., & LONG, L. Comparison of reasoning ability in two groups. *J. genet. Psychol.*, 1943, 62, 63-76.

Supplement to

CHILD DEVELOPMENT

Volume 27
June, 1956
No. 2

CHICAGO CONFERENCE ON CHILD DEVELOPMENT

1955

Contents

Child Development: An Historical Perspective
John E. Anderson 181

Some Problems of Personality Development Research
Vaughn J. Crandall 197

A Program of Research on the Dimensions and Antecedents of Emotional Dependence
Jacob L. Gewirtz 205

Some Research Problems for Developmental Psychology
Eugene S. Gollin 223

The Replacement Problem in Child Development
Harold E. Jones 237

The Need for Conceptualization in Research on Child Development
Irving E. Sigel 241

Experiments with Children on the Hypotheses of Acquired Distinctiveness and Equivalence of Cues
Charles C. Spiker 253

Psychological Development in Midwest
Herbert F. Wright 265

CHILD DEVELOPMENT: AN HISTORICAL PERSPECTIVE

JOHN E. ANDERSON
Institute of Child Welfare
University of Minnesota

As with development, history may be approached from several points of view, each of which results in a different treatment. First, there is a straightforward chronicle of events, a descriptive account of what happens, which gives minimal determination of relations other than those indicated by the time order itself. As with longitudinal studies of development, mountains of records become available from which selection occurs in both recording and retention.

A second approach is in terms of men, since particular persons acting as leaders express group tendencies and gather followers who thenceforth affect the stream of relations. Analysis in terms of leaders however identifies the social process with a few individuals. It is the history of the priests rather than that of the people. Particularly is this the case in the history of science, which, in its very essence, depends upon many intellectual workers who pursue similar lines of inquiry.

Third, there is an analysis of the trends within historical time based upon antecedent-consequent relations and the interconnections of events. This reveals the essential logic of the historical process and outlines the emergence of ideas and of intellectual movements. But this approach is, however, open to great error since a complex manifold is approached from the point of view of a theory in terms of which the material in the manifold is restructured. Very often there is no possibility of validation since history as an open end system can never be reconstituted. This problem also faces us in the study of human development.

Moreover, the symbolic system into which history must be put is quite as inadequate as is our symbolic system for representing the developmental process. Simple relations between two events, persons, variables, sequences are well handled; the complex of relations and the relations between relations make up a manifold that can probably never be represented as a totality. Our problem is complicated further by the fact that the system

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

is dead and gone; we push the dry bones of our records around in the hope that some glimmer of the life that was will emerge.

Precursors

Before discussing child development which begins as an entity in the early twenties, we may take a hasty glance at thinking about children in an earlier period. It is difficult for any of us to understand the world of one hundred years ago in which children moved through school without clear recognition of differences in capacity and growth, without understanding of the learning process, and without insight into what underlies good and bad adjustment. It was a world in which the study of children was frowned upon. G. Stanley Hall has told how difficult it was in the middle of the last century to measure the height and weight of school children because parents objected to the invasion of the child's rights. The few persons who kept records of children were scientists who used their own children as subjects. Usually they were satisfied with the production of a baby biography.

Recently I read a text widely used in 1888 for training teachers in what would now be called child development. In it there is no discussion of individual differences or of growth; only three pages are given to what would now be called learning. Memory is discussed in twelve pages without any data. Most of the book is spent on sensation and perception and on material on reasoning which now would be found mainly in texts on logic or philosophy. There are no quantitative materials, no citations of experiments, and no tables or graphs.

In the nineties several intrepid investigators such as Bolton and Gilbert tested reaction time, color discrimination and memory in children from six to sixteen. But unfortunately their tests showed little or no relation to school performance and had no practical results in placement or training. So they remained ahead of their time.

Great strides came from experiments on learning on adults. Bryan and Harter plotted their observations on learning telegraphy and discovered the learning curve. With this demonstration that progress in learning could be plotted and measured, it became clear that methods of improving instruction and learning could be discovered by experiment. This area of research was rapidly extended to children. The pioneer experiment of William James on transfer of training was followed by the important investigations of Thorndike and Woodworth in 1901, which attacked the doctrine of formal discipline and set off a great controversy and much research, most of which was done on children.

Now comes a great nodal point. In 1907, Alfred Binet, impressed by the injustice of the intuitive method of placing children in institutions then in vogue, sought an objective method for assessing the child's level that would

be independent of the examiner's judgment. He came upon the method of age scaling and demonstrated that children's performance increases rapidly with age. Soon it was clear that the measurable differences in mental level affected children's school achievement. Although differences were recognized prior to Binet, they were felt to be related to working habits, moral fiber, and willingness; in other words, to character and moral qualities over which the child had voluntary control. And let us not forget that it was a world in which many children were given severe physical punishment because they had what we now recognize as inadequate capacity.

Why is Binet's work so important? Clearly it substitutes for the onus of moral blame a measurable phenomenon within the child's resources, which is to be studied henceforth in its own right. The repercussions were tremendous. Bright children, normal children, handicapped children, retarded children, deafened children, blind children; all types and kinds of children everywhere were measured and studied in the years that followed. But, more than that, Binet made psychologists everywhere so highly aware of the relation of chronological age to mental processes that henceforth almost every investigation on children contains some analysis of the results in terms of chronological or mental age. Moreover, Binet's techniques spread widely and rapidly over the whole world. It is impossible for us, unless we lived through the period, to recapture the enthusiasm, discussion and controversy that the Binet tests started.

Then came Thorndike, who in addition to working out statistical techniques, stimulated vast research on almost everything that could be measured in school children. From this there came critical attitudes toward long-existing practices that resulted in many experiments designed to evaluate the educational process.

Here then are a number of trends in the psychological field which came together: intelligence tests, learning, transfer and interference experiments, and widespread concern with the relation between the products of education and age level. There resulted an atmosphere which made age relations, learning, and individual differences focal points for study.

Then World War I gave an opportunity to apply some of the testing techniques which had grown out of research on children to a young adult population in the military service. Although the tremendous effects of the psychological work of World War I in developing measuring techniques is beyond our scope, we may point out that hundreds of psychologists who were trained in the new measurement techniques were spread out over the entire United States after the war. This very rapid diffusion of knowledge had immediate and great consequences in providing the techniques with which a growing science attacked many human problems including those of children. A third and very striking effect of World War I with its physical, medical and psychological examinations was to present a picture of the young adult American male, which was not very pleasing. A surpris-

ing number of young men showed correctable physical defects. Illiteracy, because of inadequate educational facilities in many sections of the country, was clearly demonstrated. Many maladjusted and inadequate persons were discovered. Educational opportunities were clearly below the capacity of many who had been pupils.

The Child Development Movement

Hence, the early 1920's were times of great ferment, which extended into medicine and health as well. As the American people set out to remedy the defects in the physique and the training of their young men, new sciences and new arts relating to the physical and mental development of human beings came into being. This was the period when well-baby clinics, preventive pediatrics, school health programs, child guidance clinics, mental hygiene clinics, and demonstrations in public health put in their appearance. And among the new movements was the child development movement, which this time became a reality. Once before, in 1909, under the sponsorship of G. Stanley Hall, an attempt had been made to coordinate the activities of various scientific societies into an interdisciplinary attack upon developmental problems. But the atmosphere then was not right.

Systematic and organized study in the field did not begin until the early twenties when great interest in the growth and development of children appeared. In 1917 the Iowa Child Welfare Research Station opened with Bird T. Baldwin as Director and in 1921 it published a study of the physical growth of children which was followed a year later by a monograph on the mental growth curves of normal and superior children. In 1916 Terman began working with gifted children. In 1919 Watson's *Psychology from the Standpoint of a Behaviorist* with its strong emphasis on the importance of data on infants for the understanding of adult psychology appeared. In the early twenties, Gesell at Yale was observing the behavioral development of infants and young children, while Dearborn began the Harvard Growth Study and Freeman the Chicago Growth Study in each of which large school populations were followed from year to year with physical measurements and mental tests. Much interest in development appeared in different institutions and places. Of importance also were the continuing studies of Thorndike and his students on ability and achievement tests for children of various ages and school levels.

Spelman Funds

Into this picture there came the Laura Spelman Rockefeller Memorial with two able and active persons, Beardsley Ruml and Lawrence K. Frank, who, to meet one of the avowed purposes of the Memorial, were responsible for encouraging research in child development and parent education over the whole United States. Frank in particular played a prominent part in

the early organization of the field and maintained high personal interest and enthusiasm for the activities undertaken. Substantial funds were made available and a series of institutes whose primary function was the conduct of research in child development were established in universities. The Child Welfare Institute at Teachers College was established at Columbia in 1924, and the Institutes of Child Welfare at the University of Minnesota and at Toronto in 1925. In 1926 the Yale Psycho-Clinic was financed. In 1927 the Institute of Child Welfare at the University of California was established and substantial funds were made available to the Iowa Child Welfare Research Station. Smaller grants were also given but these were mainly concerned with the education of very young children or with parent education. In order to increase the effectiveness of these grants, funds were made available to the Child Development Committee of the Research Council for fellowships, conferences for stimulating research and for its work as a coordinating agency. The plan of the Laura Spelman Rockefeller Memorial was comprehensive. Put into action, it literally created a field that was nation-wide in its scope and thus advanced work in the area of child study by several decades.

In many of the units established under Spelman Funds, nursery schools were established in order to make young children available for observation and study. In some, funds were available for parent education programs. In several, special facilities or projects concerned with the observation of infant behavior were established. Although the primary emphasis was on the infant and young child, some projects were undertaken at later ages. As the time passed the Institutes grew with their children and more studies were undertaken on older children.

For the infant studies the most widely known center was the Yale Psycho-Clinic where Gesell developed his observational dome and made follow-up studies of many infants. But there was great interest in infant development at other places: Weiss' studies at Ohio State; the Shirley and Boyd study at Minnesota; the McGraw study which made Johnny and Jimmy famous; and the various studies of Bayley *et al.* at California.

Compared with the studies on infants which were confined to a few centers, there were many studies on children from two to five years. A whole scientific literature covering virtually every aspect of the developmental process was created within a few years by the activity of the various Institutes. Into the picture there also came a number of organizations with other sources of funds. One of these was the Merrill-Palmer School at Detroit which began in 1920, but did not become active in research until the mid-twenties. At Cornell a program in Child Development and Family Life began in 1925. The Brush Foundation at Western Reserve University was established in 1927 to carry on longitudinal studies and the Mooseheart Laboratory for Child Research was established in 1930. In 1927 the Child Research Council at Denver began operating. Later it amalgamated

OUTLINE: HISTORY OF CHILD DEVELOPMENT

	Before 1920	1920–29	1930–39	1940–On
University Units	**Individuals** Darwin 1809 — Wallin 1876 Galton 1822 — Koffka 1886 Preyer 1841 — Buhler, C. 1893 G. S. Hall 1846 — Werner 1890 Binet 1857 — Piaget 1896 J. M. Baldwin 1861 — Fraud 1856 Thorndike 1874 — Watson 1878 Stern 1871	10 Gatzert Fund became ICW, U Washington (Smith 83) 17 Iowa C W Research Station (Baldwin 75, Sears 08) 24–34 CWI, Teachers College, Columbia (Wooley 74, Jersild 02) 25 ICW, Minnesota (Anderson 93, Scammon 83, Goodenough 86) 25 ICW, Toronto (Blatz 95) 25 CD&FL, Cornell (Waring) 26 Yale Psycho-Clinic (Gesell 80) 27 ICW, California (Jones 94) 27 Child Research Council, Colorado (Washburn 95) 28 Brush Foundation, W.R.U. (Todd 85) 30 Research in Child Dev., Michigan (Olson 99) Committee Human Dev., Chicago		48 Child Study Dept., Yale
College Units		29 Fels Institute, Antioch (Sontag 01)	35 Char. Research Proj., Union (Ligon 97)	
Clinics	96 Psychological Clinic, University of Pennsylvania (Witmer 67)			
Independent Units	08 Vineland Research Laboratory (Goddard 66, Porteus 83, Doll 89) 08 Elizabeth McCormick Memorial Fund	20 Merrill-Palmer School (White 80, McGinnis 93)	30 Mooseheart Laboratory for Child Research (Reymert 83)	
Longitudinal Projects		22 Stanford Gifted Children (Terman 77) Harvard Growth Study (Dearborn 78) Chicago Growth Study (Freeman 80)	30 Neuromuscular Dev. of Infants and Johnny and Jimmy (McGraw 98)	
Handbook or Manual			31 (Murchison 87)	46 (Carmichael 98)
Foundations		18–38 Laura Spelman Rockefeller Memorial (Ruml 94, Frank 90) 18 Commonwealth Fund 24 Child Guidance Clinics (Stevenson 92)		
U. S. Government	12 United States Children's Bureau White House Conferences . 1909 ..... 1919 ..... 1929 1940			
Voluntary Lay Organizations	88 Child Study Association of New York (Gruenberg 81) 97 National Congress of Parents and Teachers 98 Child Study Program, American Association of University Women	25 Association for Family Living, Chicago (Lawton)		
Professional Organizations	1909 Hall's abortive attempt	20 Committee on Child Development, National Research Council (Woodworth 69) 26 National Association Nursery Education 26–41 National Council of Parent Education	33 Society for Research in Child Development	46 Division of Childhood and Adolescence, A.P.A.

with the University of Colorado and started a series of extensive longitudinal studies. In 1929 the Fels Institute at Antioch began its extensive program covering many aspects of child development.

In order to coordinate the efforts of various scientists and to further interchange of information about research techniques, the National Research Council revitalized its committee on Child Development in 1925. In the course of the next few years this committee held a series of conferences in which representatives from various sections of the United States and from various disciplines participated. The committee had an active secretariat with funds from the Laura Spelman Memorial. It also became an agency through which substantial sums were distributed to fellows and scholars in child development and in parent education.

Out of the activities of this committee came the Society for Research in Child Development, which was organized in 1933, and which has continued up to the present time with its various publications.

In discussing the development of the field some attention must go to the White House Conferences. The first of these called in 1909 by President Theodore Roosevelt concerned itself mainly with what the national government should do about children. Out of it came the U. S. Children's Bureau, which has been active ever since. The Conference in 1919 dramatically brought to the fore the physical and mental deficiencies revealed by the draft in World War I. The Conference in 1929 pooled substantial amounts of research from all the agencies in the United States (including much data from the Institutes), conducted various surveys on its own, and issued a substantial series of publications which are still significant. The Conferences held in 1940 and 1950, however, were much more concerned with the practical and social aspects of the problems presented by children rather than with research findings.

As an aid to visualizing this chronology there is attached an outline which locates in the time sequence various (not necessarily all) units beginning with the year 1920. In the upper left corner are the names of persons who were prominent in the field prior to or parallel to the formal units. Wherever a person is identified with one of the units his name appears immediately after along with the date of his birth.

Men

It is not my purpose to single out particular persons for comment but to generalize with reference to types of contributions made. Before 1920 the men who were interested in research on children were scientists with many widespread interests. Some, such as Galton and Darwin, were mainly interested in the whole developmental process from a biological background; some such as Preyer were physiologists who turned to embryology and the early stages of life; some such as Romanes and J. Mark Baldwin had philo-

sophical as well as biological interests and sought answers to phylogeny in ontogeny. There were no specialists in the modern sense, even though G. Stanley Hall probably comes nearest. Yet his interests were very broad. For Binet, whose influence in the public mind is closely identified with the single outcome of intelligence testing, the development of the scale was an incident in a career of many and varied contributions to psychology.

Moving into the twenties we find child development selecting its personnel from many intellectual and scientific sources. With the establishment of the Institutes what amounted to a profession was established. Even yet, however, we are in the throes of the ongoing developmental process with much controversy whether or not we are multidiscipline, interdiscipline, single discipline or an offshoot of a particular discipline. But nevertheless a considerable group of dedicated persons identify with child development as a field.

In talking with a friendly historian he made the point that in writing history one must look out for persons who exercise great effect upon a field without leaving behind the permanent records which become the material for history. Often devious searches and odd clues must be used to locate them. In the living system they are powerful determinants; in the dead system they leave few traces. We have such a person. In the McLean article on the history of Child Development (5), R. S. Woodworth is barely mentioned as a member of the Committee on Child Development. No account is given of his long chairmanship of the committee, his efforts in organizing the Society for Research in Child Development, his great services in securing recognition and status for child development in the National Research Council and the Social Research Council, and his active participation in conferences, committees, meetings and in the development of publication policies. He is the strong silent man of our history, a man famous enough outside our field, who was for us over many years a real father figure.

Trends in Content and Ideas

Growth of the Scientific Literature

One approach that can be made in evaluating the changes in time within our field is to make some counts of the scientific literature in the field. This is particularly important in view of the fact that there are the Spengler's among us such as Dennis and Barker who give the impression that we are going downhill in our knowledge of development and are farther from our goals than we were in 1925. To approach this problem I made the three following analyses.

Textbooks. The first book with the title *Child Development* that I have found is that of King (3), an associate of John Dewey, published in 1903. When one compares this book with a modern text the differences in 53 years are impressive. King's book is quite general and philosophical and,

except for an area on the activities and interests of children which refers to many of the studies of G. Stanley Hall and his students, there is a paucity of references to specific data. Compared to Thompson's *Child Psychology* (1952) with its 667 pages, its 225,000 words of text, its 1177 references, its 150 figures and its 41 tables; King has 265 pages, 45,000 words of text, 106 references, 4 figures on physical growth (as a sort of appendix) and no tables.

Handbook and Manual. A check of the *Handbook of Child Psychology* (6) and its successor the *Manual of Child Psychology* (1) makes obvious the increase in the accumulated material in the field as shown in Table 1 which gives the number of words for the book as a whole and for the only two chapters (Anderson, McCarthy) which have covered the same content in all four editions. Obviously such figures only approximate the trend since editorial considerations determine in some degree the amount of material presented.

TABLE 1

WORDS* IN *HANDBOOK OF CHILD PSYCHOLOGY* AND *MANUAL OF CHILD PSYCHOLOGY*

			Entire Book	*Anderson Methods*	*McCarthy Language*
Handbook	1st Ed.	1931 ...	275,000	11,000	14,500
Handbook	2nd Ed.	1933 ...	405,000	13,500	18,000
Manual	1st Ed.	1946 ...	525,000	23,000	56,000
Manual	2nd Ed.	1954 ...	650,000	32,500	73,000

* Exclusive of bibliographies.

References in Psychological Abstracts. Recently I have had the references entered in the yearly indexes of *Psychological Abstracts* counted under the headings "Child," "Infant," "Adolescence," and "Family" from the first volume in 1927 up to the present. The results are presented in Table 2 by seven year periods.

Far from showing a decline in interest in children or research upon them, these figures show an increase in the absolute number of references and a fairly constant percentage relation of about 7 per cent which indicates that publication in the child field varies with that in the total field. "Infant," "Adolescence," and "Family" have a small number of references with some increase in absolute numbers and fairly constant percentages.

While these different analyses concern only the quantity and not the quality of our intellectual and scholarly production, they certainly do not indicate that we are worse off than we were in an earlier period. While the Golden Age in which the Institutes were being established and which is

TABLE 2

MEAN NUMBER OF INDEX REFERENCES AND PERCENTAGE OF TOTAL PER YEAR FOR VARIOUS TOPICS IN *PSYCHOLOGICAL ABSTRACTS* FOR SEVEN YEAR PERIODS FROM 1927 TO 1954

Period	*Child* Mean per yr.	%	*Infant* Mean per yr.	%	*Adolescence* Mean per yr.	%	*Family* Mean per yr.	%	*Total Abstracts* Mean per yr.
1927–1933	325	6.7	8	0.2	38	0.8	29	0.6	4704
1934–1940	437	7.0	45	0.7	53	0.8	30	0.5	6270
1941–1947	386	8.5	27	0.6	44	1.0	30	0.6	4555
1948–1953	501	6.9	36	0.5	72	1.0	62	0.9	7362
1927–1953	412	7.2	29	0.5	52	0.9	38	0.7	5723

reflected in the publications for the late twenties and early thirties is over, it is my opinion that there has been an increase in interest in research on children throughout the entire field of psychology (the field with which I am most familiar) and that the Institutes are no longer quite as much in the center of the picture.

Approaches to the Study of Growth and Development

In considering the content of child development we need some conception of the nature of the growing human person or a theoretical model against which to orient our studies. The growing person is an ongoing manifold moving through time with a multidimensional head much as a tunnel-making machine moves under a river. Viewed from the moving front or the present moment of time this consists of many irregularly shaped structures of various sizes with a complex series of interrelations. At the boundaries there are complex relations to the external world which involve a continuous intake and outgo on both a physiological and a psychological level, an interchange which maintains a dynamic or ongoing equilibrium.

Viewed as a whole the system is asymmetrical and open-end. Its basic organization is hierarchical in the sense that as it moves forward in time new properties and qualities emerge at successive levels of integration. These need study and analyses in and for themselves. An analysis at one level, however adequate, cannot explain the properties of a higher level.

In other words we cannot solve the problems of children's language by analyzing their pronunciation of single letters or syllables, but must study his behavior for words and sentences. The social behavior of the child must be studied in its own right by its own methods since there are limits to the extent to which it can be predicted from knowledge of the individual children in the group.

At once this model denies the possibilities of explaining human behavior on the basis of physical models and makes the psychological model analagous to those of theoretical biology. In exploring the possibilities of this model we are concerned with the study of its total shape or form as well as the shape, form, and interrelations of the parts. Moreover, analysis of time relations and of transformations in time become crucial.

Many of the apparent contradictions in behavior and development theory result from the fact that we study different aspects of this exceedingly complex manifold as it undergoes transformation in time. To the complexities of the unfolding system are added the inadequacies of our symbolic system which, whether verbal, pictorial or mathematical, has distinct limitations (9).

Using our model as a base we may have a look at the accumulated scientific literature in our field which we can divide into eight broad classes, each of which could be treated in considerable detail.

1. *Space-time form studies.* These are concerned with the properties and interrelations of systems within the whole model as it evolves in time. After the comparable biological science which studies the transformations in physical growth, this area is a kind of psychological morphogenesis. Either successive cross section slices or a combination of many longitudinal slices may be used to create the picture. Although such studies result in a wealth of information about the parts of the model, because of their rarity emphasis could well go not so much to the slices or factors as such as to the interrelations between variables and the emergence of higher levels of behavior. This is the most involved type of study and, except for the longitudinal studies at California, we have hardly made a dent in this area.

2. *Longitudinal slices.* There are a number of studies in which some particular or several aspects of development are followed in the same subjects over a period of time. The distinction between these slices and the studies under the previous heading is one of extent. Related to both are the follow-up studies, now being set up in many places in which records obtained much earlier are used as a base to study continuities or discontinuities. The Terman studies of intelligence are an example in point.

3. *Cross section studies at successive levels.* These analyze segments or relations over substantial periods of time by selecting particular aspects of development at successive age levels on comparable samples of the population in order to plot or lay out the course of development. These studies become concerned with the constancy or change problem, and with the analysis of the underlying factors which produce constancy or change. Problems of definition and classification of the ongoing phenomena at successive age levels also arise. There is a great bulk of studies of this type in many areas. At present we can write a respectable picture based upon substantial and reasonably complete research of the development of the child from birth to maturity in many psychological areas, particularly intel-

ligence and language. There are, however, some gaps in our knowledge of perceptual, motor, concept, personality and social development.

4. *Cross section studies at one level.* Many persons are interested in children at a particular level. Some are interested in infants, some in young children, some in adolescents, etc., and think only in terms of the particular level. In the design of experiments this approach leads to cross section slices taken on various samples without forward or backward reference in the developmental sequence. There are literally thousands of such studies essentially at a descriptive level which are difficult to assess or to fit into a comprehensive picture of development. It seems to me that the day of such studies is pretty well over, and that we should urge scientists who work with children to study at least three separate age levels in order to determine the relation of the phenomena in question to age. If the primary concern of the developmentalist is with changes in time, chronological age or some derivative of it must become the independent variable in our studies.

5. *Short-period function modifications.* These introduce specific stimulation compressed into a relatively short time on a sample of children. They are concerned with the mechanics of stimulation and response, or the manner in which patterns of behavior develop. One method of instruction is compared with another or a learning or problem solving situation is set up. In terms of design these studies select groups of children at some level, introduce stimulation for a time and then measure response. Learning curves are plotted and analysis is made. The experimenter loses interest in the child as soon as he reaches one or ten perfect trials or attains a defined level of proficiency and is not interested in what the child does with what is learned, nor how the content is incorporated into the total system of behavior. Rarely are such problems carried out on children at several age levels, as the investigator is only incidentally concerned with the developmental ranges through which the principles derived operate. Although the literature on perception and learning in children is substantial, it is much more concerned with simple than with complex processes.

6. *Genetic determinism.* Many studies are concerned with the unfolding or the relations of hereditary factors. Resemblance coefficients between parents and children and between siblings are commonly used. We need more and better studies of the contribution of genetic variance to the total variance in behavior, and particularly studies in which children are examined at the same ages as their parents were, in order to minimize differences due to developmental level.

7. *Environmental impingement and determinism.* These study the positive effects upon the whole and its parts, of the environment as a whole or in part along with the negative effects of deprivations of various types. Here again we move into a vast array of studies many of which are of very limited scope and offer little or no possibility of effective integration into the literature. Many are poorly designed, done on very limited age samples,

fail to isolate significant factors, and, by making superficial analyses of the relations involved, often create artifacts. Many fail to test for the permanence of the effects. In the deprivation studies which are so popular now, there is often lack of good controls, failure to recognize the basic capacity of the organism for self repair when returned to a good environment, and a general lack of sophistication about scientific method. But studies of the responses of children in their habitat, of the Barker and Wright type, offer new methods which can be widely used and suggest an entropy theory of stimulation rather than the common nickel-in-the-slot one.

8. *Construct determinism.* These studies reify a quality or attribute of the system and study the outcomes of the construct so derived. One phenomenon, usually motivational, is generalized and its effects studied. Thus we have frustration, anxiety, rejection reified into factors which thereafter determine behavior. Similarly we reify the "only child" status and ordinal position into determiners. The reified characteristic becomes a sort of product which once deposited in the ongoing system operates virtually indefinitely without modification. Here we need to find the iterative mechanisms that keep constructs, once we are sure of their validity, alive and active in the developing system. Many of the studies in this area are made on very small samples of cases at limited age ranges, but with wide generalizations to populations of all ages and types. In addition to substantial samples emphasis should go to observations on at least three age levels in order to discover what happens with age.

Areas of Study

As Woodger (9) has pointed out, the first approach to a growing organism involves cataloging and describing the many structures, functions and behaviors of the organism as it moves through time. This cataloging function which is typical of the early stages of life sciences is essential in order to locate the problems and to develop the hypotheses that make more complex generalizations possible. In the period from 1920 to 1930 workers in child development and related fields were mapping out details of physical and mental growth and particularly those from birth to six years. By fostering infant laboratories and nursery schools within the research organizations, Spelman financing directed the attention of many workers to the infant and the young child. There was much concern with the standardization of measuring instruments, such as intelligence tests for infants and young children, and much interest in language, problem solving and social behavior. Perhaps the most original developments of this era lay in the field of social behavior as the extraordinarily large number of investigations summarized by Murphy, Murphy and Newcomb (7) indicates. The sociometric techniques were also emerging. There were scattered investigations on the learning of young children.

At the White House Conference in 1929-30 a committee selected a stratified sample of young children over the whole United States for study of the relation of socio-economic status to standards, facilities and methods of care. There followed many studies of effect of social class and social status upon the development of the child with very clear demonstrations that socio-economic or cultural status is a significant determiner of variance in the goals and ideals of the parents, the methods of child training and the facilities made available to them for learning and adjustment.

Another trend resulted from the studies in the late twenties on diet selection which led to an almost complete reversal of practices with regard to schedules in the training of infants and to the movement toward self-demand rather than forcing in the socialization of the infant and young child.

Lewin and his students in the late thirties performed the experiments on autocratic and democratic atmospheres which resulted in substantial research and new conceptions of the relation of the child to its environment. The principles were later extended upward into adult behavior and revitalized areas in social psychology. In our field came the studies of home and family atmospheres with much improvement in our descriptions and analysis of the family environment and of parental behavior. Meanwhile, however, the cataloging function continued and the first results of the longitudinal studies began to appear in substantial degree. In the early 1940's psychoanalytic concepts, particularly relating to the effect of deprivation and institutionalization or the relation of early handling of the child and subsequent development, began to appear and much concern, both experimentally and theoretically, for the last decade and a half has centered in that area.

Then came World War II which temporarily diverted many of the workers in the field directly to war activities and others to such practical matters as the care of children during the emergency and technical assistance in the war or defense efforts. The organizations devoted to research on children kept going on in one way or another, but with reduced loads and sometimes with change in emphasis in their research program. So great were the disturbances that it took five or more years to return to an even keel.

To some degree just prior to the War and strikingly thereafter came the marked developments in projective techniques which opened new areas of personality measurement for both children and adults. These are still centers of controversy as well as of active research. After the war there appeared the habitat studies and techniques of Barker and Wright with their complete description of the child-environment interaction and the conceptualization of methods for analyzing the environment.

Recently there is a striking new development represented by the follow-up studies based either upon a longitudinal sample earlier followed for

many years or the follow-up as adults of the many children on whom the Institutes obtained records in the late twenties and thirties. In some degree the research of the Institutes has grown older with their children. Some effort is going into late adolescence and early adult life. Whether with these developments as much research energy can go to young children as in the early days remains to be seen.

Conclusion

In my summary of the developing content within the field I have used three approaches: first, quantitative estimates of the growth of the scientific literature; second, an analysis of our different approaches to the model of the ongoing manifold changing in time which we know as the growing human being; and, third, a brief summary of developments in content and methods by areas in the past three decades. None of these gives us any reason for discouragement or apology. A scientific literature of wide scope has been created where there was almost none before.

In large part this is the product of persons and institutions who became interested in human development in the twenties and who with their students became the professionals in the field. But also because development has acquired great significance in theories of behavior and adjustment, much work has been undertaken outside our channels. The briefest search of the extant literature will show such interests widely scattered among psychologists, sociologists and anthropologists. And quite recently a surprising amount of interest in the young of the species has appeared among the students of animal behavior.

But even more impressive than the extent of the literature are the new methods of approaching the behavior of the child, of analyzing the stimulation of the environment, and of studying the modification of behavior. These give us a surprisingly rich armament with which to move forward. And, as with any science in a period of rapid progress, new problems continue to emerge.

Finally I feel that we are essentially a biological science, studying an open-end living system of extraordinary complexity. We cannot hope for the extreme condensation of modern physics; the equation of a living organism, if such could be written, would fill many blackboards. We have done much cataloguing and describing and must continue to do so. But this is not enough; we must move on to first order generalizations that are close to our data and then on to those of the second, third and n orders. Whenever we can substitute experimentation for description we will move ahead. But let us not assume that a theoretical model erected for a limited portion of the relational system is the equivalent of the model for the whole. We need better methodology; we need better theory; we need more adequate sym-

bolic systems if we are to picture and conceptualize what has been often described as the most interesting phenomenon of life: the fertilized cell becoming the complex manifold that we know as the mature human being.

References

1. Carmichael, L. (Ed.) *Manual of child psychology.* New York: Wiley, 1946 (1st Ed., 1068 p.), 1954 (2nd Ed., 1295 p.).
2. Dennis, W. Historical beginnings of child psychology. *Psychol. Bull.,* 1949, 46, 224-235.
3. King, I. *The psychology of child development.* Chicago: Univer. of Chicago Press, 1903.
4. Kirkpatrick, E. A. Child study. *Pop. Sci. Mon.,* 1910, Nov.
5. McLean, Dorothy. Child development: A generation of research. *Child Develpm.,* 1954, 25, 3-8.
6. Murchison, C. (Ed.) *Handbook of child psychology.* Worcester, Mass.: Clark Univer. Press, 1931 (1st Ed., 711 p.), 1933 (2nd Rev. Ed., 956 p.).
7. Murphy, G., Murphy, Lois B., & Newcomb, T. M. *Experimental social psychology.* New York: Harper, 1931 (Rev. Ed., 1937).
8. Thompson, G. *Child psychology.* Boston: Houghton Mifflin, 1952.
9. Woodger, J. H. *Biology and language.* Cambridge: Cambridge Univer. Press, 1952.
10. *Psychological Abstracts,* Vols. 1-29. Washington, D.C.: Amer. Psychol. Ass., 1927-1954.

SOME PROBLEMS OF PERSONALITY DEVELOPMENT RESEARCH

VAUGHN J. CRANDALL

Fels Research Institute for the Study of Human Development

The content and methodology of personality development research, like those of the total field of child development research, vary tremendously. Past personality development studies have ranged from empirical normative surveys to theory-oriented hypothetical-deductive ones, and from naturalistic observational investigations to controlled laboratory experiments. No attempt will be made in the present paper to survey all, or even most, of the current theoretical and methodological problems in this area of child development research. Rather, a limited number of these problems will be selected for discussion, with special emphasis on their implications for future research.

A number of personality development studies in the recent past have employed projective techniques as research tools. While some of these techniques have been very useful for studying certain problems, projective techniques have often, unfortunately, been employed in an uncritical fashion. Too frequently the tail has wagged the dog, with research problems being determined by available projective methods rather than by specific theoretical questions. It is the belief of the writer that this state of affairs can be, and probably will be, eradicated as research in this area becomes less empirical and more theoretically oriented. Ideally, in theory-oriented research, hypotheses under consideration should dictate behaviors to be studied and thus, indirectly, determine the choice of assessment techniques to be used. Existing projective techniques may or may not be appropriate for such investigations. In sheer number there is no shortage of potentially useful projective tests; the growth of these assessment methods has been almost weed-like. However, since many of these techniques were originally developed with little regard for existing personality theories, few may be found useful in future personality development research in their present form. New techniques will have to be developed and/or old ones modified. As the writer suggested in a recent symposium on the use of projective tech-

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

niques for the assessment of normal personality development (5), several existing projective media appear particularly amenable to modification for theory-oriented personality studies. For example, specially-constructed sets of TAT-like pictures can be employed to test a variety of theoretical hypotheses. Blum's use of his Blacky Pictures to test selected psychoanalytic hypotheses is an example of this (1), as is the writer's employment of specially-constructed pictures to test Social Learning Theory (3, 4, 13). Incomplete sentences and stories, too, can easily be devised to elicit behavior relevant to specific theoretical hypotheses. Semistructured imaginative play situations can be used in a similar fashion. When existing projective media cannot be modified to fit research needs, new techniques will have to be devised. In theory-oriented research, however, the construction of these will be dictated by the problems under consideration rather than a desire to develop or use a new projective technique *per se* as has frequently happened in past research.

A second problem inherent in the use of projective techniques in personality development research is that of the categorization of responses. Since projective techniques are designed to elicit relatively "free," unstructured responses, these responses must be subsequently categorized by the researcher in some fashion. In past research, projective test responses have usually been analyzed according to a hodge-podge of empirically-derived and theoretically-unrelated categories. "Sign approach" methods for analyzing Rorschach protocols typify this approach. Such empirically-derived systems can seldom be used to categorize responses evoked by techniques other than those for which they were originally devised. As a result, data analyzed in such a manner usually lack generality. These data can neither be directly compared with data gathered by other personality assessment methods, nor can they be readily related to existing personality theories. Future personality research should include methodological studies concerned with the development of theory-oriented categorizations of responses applicable to data elicited by a variety of personality assessment methods. Ideally, such systems of categorization should be equally suitable for objective, subjective and projective responses when possible. Examples of existing methods of analysis which might be used in this fashion are Raimy's PNAv method developed within the framework of self-concept theory (12), Dollard and Mowrer's Discomfort-Relief Quotient influenced by Hullian learning theory (8), and the writer's Social Reinforcement Index developed within the context of Social Learning Theory (7, 13).

At the present time our knowledge of the areas and levels of behavior elicited by various methods of personality assessment is far from complete. Future personality development research should be facilitated to the degree that *cartographic studies* of existing assessment methods are a part of this research. These investigations could systematically map out areas, levels and frequencies of behaviors elicited from children of various ages by each

assessment technique. Ultimately, this research should answer, among others, the following questions: (*a*) To what degree does a given technique cover the same behavioral areas with children of various ages? (*b*) At what level of expression is this behavior elicited? (*c*) Does the method produce an adequate sampling of this behavior to insure stability of measurement? (*d*) Is the range of this behavior great enough to allow sufficient intersubject and intrasubject variability? Such information should be helpful, not only to evaluate the usefulness of existing assessment techniques as research tools in personality development research, but also to determine areas where new methods must be devised. Cartographic information of this nature should ultimately open new avenues of research.

When suitable methods are available, for example, intraindividual covariational studies of children's objective, subjective and projective behavior will be possible. Such research will, of course, require assessment techniques covering comparable areas of behavior at these three levels of expression. When techniques are available, a variety of developmental studies are feasible. For example, while a number of recent investigations have dealt with the acquisition of social needs in children, there have been no studies, to the writer's knowledge, on changes in levels of expression of these needs as children mature. Such changes should be an integral part of a child's socialization. He comes to learn that the overt expression of certain needs is socially acceptable while a similar expression of other needs is not. Studies of adult subjects have indicated that correlations between objective and projective expressions of socially acceptable needs differ predictively from those of socially unacceptable ones. Future personality development research could study changes in such correlations in children's behavior with age, as well as factors producing these changes. Other investigations of a similar nature could be conducted in respect to the development of ego defenses in normal children.

The discussion so far has made no distinction between longitudinal and cross sectional investigations of personality development. Large-scale, long-term longitudinal projects such as those currently underway at several child research centers throughout the country are faced with one major problem not found in cross sectional research. This is due to the fact that, by the time large-scale longitudinal programs are organized and data are collected for several years, original concepts and methods of measurement are often obsolete. Researchers conducting longitudinal projects are continually faced with the decision of whether to introduce new concepts and methods in the study as it progresses, thereby sacrificing comparability of data collected from year to year, or to retain original methods and concepts when better ones are available. Obviously, there is no clear-cut solution to this problem.

One partial solution is what might be called *serial longitudinal research*. Long-term research projects following the development of children from

birth to adulthood could be divided into a series of smaller studies, each covering consecutive two- or three-year periods. This design would allow continuity of subjects for the total investigation. In addition, some of the more durable concepts and methods could be retained throughout. When necessary, however, new methods and concepts could be introduced whenever a new short-term study was started. In this kind of program, each of the smaller studies could stand by itself as an independent investigation directed toward one or more problems germane to the period of personality development under consideration. In addition, data gathered from each of these studies could be used to make predictions concerning the subjects' behavior during later periods of personality development. Serial predictions of this kind, when checked against predicted-to behavior, should provide useful information on factors producing continuity or change in children's personality development as they mature.

Some of the studies might profitably focus on *transitional phases of social learning*. In our culture, while there are many periods of a child's life when his environment is relatively stable, there are also times when he is faced with rather drastic changes in his daily experiences. At these times, he encounters situations requiring the rapid learning of new social behaviors and the discarding of old responses when they are no longer appropriate. An example of one of these phases is the beginning of nursery school. The child is suddenly required to leave the stable, secure and familiar surroundings of home and mother for new, and oftimes threatening, experiences with strange children and nursery school teachers. This and other transitional periods should be particularly fruitful ones to investigate with intensive, short-term longitudinal studies; at these times the effects of specifiable environmental changes on social learning and personality development should be highlighted, affording the developmental researcher valuable opportunities for observing and measuring those relationships.

So far, this paper has discussed several methodological problems of personality development research. The remainder of the paper will be concerned with one basic theoretical problem, the generalization of social reinforcements. To the degree that personality development is a product of social learning experiences, the question of how social reinforcements generalize is an important one for personality development research. How do the effects of a given social experience determine a child's expectations and behaviors in subsequent situations? When a child experiences social rewards or punishments for his behavior in one situation, it is assumed that the effects of these reinforcements generalize to other situations in proportion to the degree of similarity between the original and subsequent situations. At the present time, however, we have little knowledge of factors determining psychological similarity in social situations. There is a paucity of information on the types of generalization operative in children's ongoing social behavior as well as the dimensions determining this generalization.

Several types of generalization may play a role in social learning and personality development. First, social reinforcement generalization may sometimes be a function of the *physical similarity of the stimuli* in the situation. Classical conditioning studies, e.g., those of Hovland (9, 10), suggest that this type of generalization is a reliable phenomenon in the laboratory. Whether it is important in children's social learning, however, must be determined in future personality development research.

A second kind of generalization which may influence children's social learning is that mediated through *similarities in conceptual-language structure*. Certain studies on semantic generalization and concept formation are relevant here. So too are some investigations on the development of prejudice. Few of these studies, however, have been primarily concerned with this type of generalization in social learning. In this respect, conceptual-language structure may be one important factor in personality development by functioning as a kind of perceptual-cognitive filter determining the effects of social reinforcements and their generalization. Until the problem is investigated directly, however, its role in social reinforcement generalization is an open question.

Social reinforcement generalization may also be a product of the *functional similarity of social reinforcing agents*. This has been suggested in psychoanalytic theory in such concepts as "father substitute" and the like. Unfortunately, most of these concepts are vague. They neither specify the particular characteristics of reinforcing agents producing reinforcement generalization, nor situational conditions conducive to its occurrence. Future personality development research can gain leverage on this problem in at least two ways. First, preliminary studies can be conducted determining the characteristics which children of various ages impute to specific reinforcing agents in their lives. Projective materials may be used in this research or actual persons may be employed as stimuli. The writer has found the latter useful with college students (6), but has not yet employed it with young children. Second, experimental studies can be devised in which the characteristics and functions of reinforcing agents are systematically varied, and resultant changes in children's behavior measured.

Social reinforcement generalization may also be mediated through the *similarity of an individual's social goals or needs*. Social behavior is characteristically goal-directed. A child, as the result of past social learning, develops certain needs and attempts to satisfy these through the attainment of corresponding social goals. When this goal-directed behavior is rewarded or punished, the child's expectations for future social reinforcements in respect to this behavior are altered. This experience may also alter the child's expectations and behaviors in respect to other similar social goals. The amount of this generalization would be expected to decrease with decreased similarity between the goal of the original behavior and other goals. Thus, for example, if a child is punished or rewarded while attempting to gain aca-

demic recognition, the effects of this reinforcement should generalize more to his behaviors directed toward other similar goals, such as social recognition or physical skills recognition, than to his behaviors concerned with dissimilar goals, e.g., the attainment of affection. Original research by the writer (4) and subsequent investigations by Jessor (11) and Chance (2) have demonstrated that such a process does occur in some social learning situations. To date, however, research on this phenomenon has been limited to adult subjects. Future child development research might include both studies dealing with this type of generalization in children's social behavior *per se* and investigations determining the role of this process in the original development of children's needs and social goals.

The four types of generalization which have been discussed do not, of course, exhaust the ways in which effects of social reinforcements may generalize in children's social learning experiences. Other types of generalization may also exist. Some of these may be identified in future animal research or classical learning and conditioning experiments. Still others may be elaborated in studies of complex social behavior. Regardless of how they are originally identified, however, it will be the responsibility of future personality development research to determine ultimately their respective roles in children's social learning and personality development.

References

1. Blum, G. S. A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monogr.*, 1949, 39, 3-99.
2. Chance, June. Generalization of expectancies as a function of need-relatedness. Unpublished doctor's dissertation, Ohio State Univer., 1952.
3. Crandall, V. Induced frustration and punishment-reward expectancy in thematic apperception stories. *J. consult. Psychol.*, 1951, 15, 400-404.
4. Crandall, V. An investigation of the specificity of reinforcement of induced frustration. *J. soc. Psychol.*, 1955, 41, 311-318.
5. Crandall, V. Projective techniques in child development research. Paper read at Amer. Psychol. Ass., San Francisco, Sept., 1955.
6. Crandall, V., & Bellugi, Ursula. Some relationships of inter-personal and intra-personal conceptualizations to personal-social adjustment. *J. Pers.*, 1954, 23, 224-232.
7. Crandall, V., & Bellugi, Ursula. Rating personal adjustment through an analysis of social reinforcements. *J. consult. Psychol.*, 1956, 20, 49-52.

8. Dollard, J., & Mowrer, O. H. A method of measuring tension in written documents. *J. abnorm. soc. Psychol.*, 1947, 42, 3-32.

9. Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. genet. Psychol.*, 1937, 17, 125-148.

10. Hovland, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, 51, 279-291.

11. Jessor, R. The generalization of expectancies. *J. abnorm. soc. Psychol.*, 1954, 49, 196-200.

12. Raimy, V. C. Self-reference in counseling interviews. *J. consult. Psychol.*, 1948, 12, 153-163.

13. Rotter, J. B. *Social learning and clinical psychology.* New York: Prentice-Hall, 1954.

A PROGRAM OF RESEARCH ON THE DIMENSIONS AND ANTECEDENTS OF EMOTIONAL DEPENDENCE

JACOB L. GEWIRTZ[1]
The University of Chicago

The problem set for this paper is to outline a program of research which focuses upon an assumed continuity between early childhood conditions and those consequent social behaviors of children which typically have been labeled "emotional dependence." In particular, emphasis is placed upon behaviors which are employed apparently to gain such positive social responses from other persons as those involving attention, approval, affection, reassurance, and nearness, and upon the early conditions under which they develop. While it is probable that knowledge of the processes underlying emotional dependence will be critical to the understanding of a substantial portion of the social behavior of children and adults, very little is known either about the dimensionality or the antecedents of that behavior class.

The principal focus of this paper is upon child behavior; however, the major problems raised may be considered a subclass of general behavior problems and hence continuous with questions characteristically raised about adaptive behaviors exhibited in simple settings both by infrahumans and humans. Since the theories of social development which do not have their origins in such basic considerations of learning appear to approach the earliest conditions for the development of dependence in a less articulate manner, selected descriptive concepts and postulates which stem from work on general behavior problems are emphasized in this paper. Relatively little emphasis is placed upon concepts at other than simple levels of analysis (e.g., personality characteristics of the mother, her acceptance of the child, social class membership), which appear to be employed in many theories and researches on children's dependence, and index conditions which would qualify early dependence learning. However, it is expected that the reader will bring to mind such concepts, theories and research findings as they are required.

A systematic ordering scheme would appear requisite to any research program of intended generality and significance. It is preferable that such

[1] The writer is grateful to Mr. D. M. Baer for his valuable aid in the preparation of this paper.

a scheme be susceptible of simple, operational statement and include a neutral, parsimonious set of terms, yet be capable of flexible application to a variety of data, procedures and samples of subjects (*Ss*). Such a system could generate hypotheses, as well as articulate with other systems of language and theory. These are rather general criteria. Additional special criteria are suggested by the problem at hand: In particular, it would seem fruitful to be able to analyze the apparent complexities of emotional dependence into rather skeletal terms so that they may be related in a simple manner to the fundamental mechanisms operative in the developing child. Thus the ordering system required must at once be capable of reducing to simple terms the complexities of the behaviors which characterize emotional dependence and of positing basic acquisition processes which may be applied to order the multiplicity of environmental conditions to which children are typically subject. And if, in addition, the properties of this ordering system were quite readily reconciled with extant concepts and theories, considerable advantage would accrue.

The basic definitions and principles in the sections which follow are intended to meet the criteria enumerated above, and hence prove useful in researches of the type required; and, further, they can serve to develop some of the background issues, both of terminology and substance, to provide the frame of reference in terms of which the special problems involved in such a program of research can be considered. It is hoped that the reader will bear with us in the consideration of the familiar concepts which will be presented in the sections which follow. The concepts and assumptions employed are drawn heavily from those of the descriptive behaviorists (6, 8). The basic paradigm employed, it will be seen, is that of *instrumental* or *operant conditioning*.

Basic Descriptive Concepts

The Operant

An instrumental or operant *response* is any "voluntary" (non-reflex) response which is originally emitted more or less independently of identifiable stimuli and which can manipulate the environment to produce changes which are of consequence for the organism. While this definition embraces much of human behavior, in this paper interest will center upon such responses as those employed for or maintained by the attention, approval, affection, proximity, caresses, and similar characteristics of another social being. While a particular response may be "spontaneous" in its initial occurrence (e.g., the gross movements of the infant), as a consequence of reinforcement it very soon can become associated with particular stimuli, such that it becomes more likely to occur in their presence. A particular stimulus which "sets the occasion" for a particular response has been termed a discriminative *stimulus* (S^D) or cue, and the response involved has been

termed a disciminated operant. It is assumed that discriminative stimuli constitute the major portion of the controlling environment for an organism. With reference to responses characterizing emotional dependence, discriminative stimuli are typically afforded by people, and thus may be termed *social.*

The Reinforcer

Any stimulus event which follows a response and affects certain of its aspects, e.g., its rate of emission, amplitude, or latency, is termed a *reinforcing stimulus* or *reinforcer.* Reinforcing stimuli are of two basic types: *positive* reinforcers strengthen responses when presented; and *negative* reinforcers strengthen responses when removed. Many of the stimulus events which afford such control over responses are supplied by people (e.g., affection), and may be termed *social* reinforcers. The concepts of response, discriminative stimulus, and reinforcer appear to allow the ordering of the phenomena covered by the rubric emotional dependence. Social behavior in general has the appearance of being directed; that is, it may be analyzed as employed to attain some happening or stimulus "commodity" from other people, like attention or approval. The concept social reinforcer directs attention to this characteristic, insofar as it focuses upon the stimulus events which yield control of responses when their presentation is made contingent upon them.

Drive

However, it is frequently observed that even in the presence of discriminative stimuli setting the occasion upon which a response shall be reinforced (by the contingent supply or removal of a reinforcer), the response is not emitted, or, if it is, the contingency of the reinforcer upon the response yields no control over it. This is tantamount to stating that on such an occasion the reinforcer is not operative as such, i.e., that there exists variance in its efficacy. This variance is generally attributed to a condition of *drive* holding for the organism at the time in question, particularly when it is demonstrated that the stimulus event in question acquires reinforcing power for the organism only subsequent to his recent deprivation of it, and loses this power subsequent to his satiation by it. In general, such stimulus events appear to possess this characteristic innately, e.g., food and water. Other stimulus events which possibly meet the same criterion may be some whose reinforcing value was acquired through a history of conditioning.

At the behavioral level, the criterion for what is and what is not a drive operation is arbitrary, i.e., one of definition and convention. In general, treatment variables are classified primarily in terms of their effects in behavior. Hence, for example, the drive label may be reserved for all nonconditioning operations which affect the efficacy of reinforcers, or it may be restricted to but a few operations which appear to have similar effects

and, perhaps also, which are judged similar according to other criteria. So also, conditioning operations in general may be classified in a number of ways. The particular classification system one employs would depend on the behavioral effects he has focused upon, the data accruing from earlier work in the problem area, labeling conventions in that area, and the conceptual system (if any) within which framework he has been working. For much work, labeling operations and advancing hypotheses are indistinguishable processes.

While the concept of drive has a variegated and somewhat confused tradition, in this paper it will be taken generally to represent almost *any* operation, apart from a standard conditioning procedure,[2] which can be shown to affect changes in response output rates. Hence, where reinforcing stimuli are identified, a drive operation is taken to be one which affects the relative efficacy of a number of reinforcers. Since for the purposes of this paper it is convenient to emphasize reinforcers rather than responses, the latter working definition of a drive operation will be employed. Where reinforcing stimuli appear to show no variance in their effectiveness, the concept of drive is unnecessary, although the identity of such reinforcers might be of significance for a variety of problems. Emotional stimuli which produce "states" like those termed frustration and conflict also may affect the relative efficacy of reinforcers. Hence the conditions producing such stimuli may be termed drive operations as well, for present purposes.

Conditioned Reinforcers

A stimulus event which is innately reinforcing is termed a *primary reinforcer* (S^R), and if its deprivation affects its reinforcing properties, then such a deprivation is ordinarily taken to define a state of *primary drive*. A discriminative stimulus which acquires a reinforcing character only subsequent to a history of pairing with a primary reinforcer, typically (but not necessarily) with the organism under a state of deprivation for that reinforcer, is labeled a *conditioned or secondary reinforcer* (S^r). It is usually assumed that the effectiveness of a conditioned reinforcer varies with the deprivation appropriate to the primary reinforcer to which is has been conditioned.

There exists the possibility of pairing a conditioned reinforcer with more than one primary reinforcer. This stimulus event has been labeled a *generalized reinforcer* by Skinner (8), and is assumed to be effective subsequently under any one of the deprivation conditions appropriate to the primary reinforcers to which it has been conditioned. Since at least one

[2] The criterion used here is that such an operation may be instituted without requiring more than at most a very few further trials to produce its effect. Such operations typically can be instituted at any moment for an organism, although the operation itself may not be independent of a history of conditioning (e.g., possibly the deprivation of a conditioned reinforcer, or breaking into an existing chain of responses in a manner which prevents completion of the chain, as in frustration).

appropriate state of deprivation is likely to prevail for an organism at any given time, the generalized reinforcer would be effective at almost any time. Further, it appears reasonable to expect that typical child-rearing practices may condition certain discriminative stimuli to a certain number of primary reinforcers under their appropriate deprivation states; hence this construct may have much relevance to emotional dependence learning.

Conditioned Drives

In the past, the apparent potency of a number of stimulus events which may meet this description of generalized reinforcer has supported the usage of terms like *social, secondary* or *conditioned drives,* e.g., for attention, approval, or the company of people. Under the present definitional scheme this usage would be in error if the effectiveness of these reinforcers could be shown to be under the control of all pertinent primary deprivations; then the phenomena properly would be referred to the characteristics of generalized reinforcers. However, it appears reasonable to propose that if the deprivation of a conditioned or a generalized reinforcer would increase its own effectiveness as a reinforcer, and if its satiation would reverse that effect, then such a deprivation may be defined as a *conditioned drive* operation. It is important that *this* definition turn essentially upon this notion of deprivation-satiation operations for the conditioned reinforcer in question; for it is usually observed that the effectiveness of a conditioned reinforcer can be referred to the organism's deprivation of the particular primary reinforcer involved in the conditioning.

If it can be demonstrated that this criterion for conditioned drive is met when children are deprived of particular social reinforcers, then a construct of great potential power for ordering social behavior would result. Since under typical social conditions many social reinforcers are not available to an individual for moderate or long periods of time, the apparent prevailing potency of these reinforcers could be assigned to their empirically accessible deprivation properties. Further, the fact that the deprivation of some social reinforcers might produce drive effects while the deprivation of others might not, would imply that certain qualitative differences in socialization procedures probably maintain. The membership of the events in each of these two classes of social stimuli would provide significant clues to these differences.

The convention about drive maintained in this paper does not require that the responsiveness of the potency of a conditioned reinforcer to its own deprivation be the only criterion defining conditioned drive. As stated earlier, drive is invoked here as a construct only to account for variance in the efficacy of a reinforcer. There may be a variety of operations distinct from standard conditioning operations which affect the potency of reinforcers in a similar manner, and drive as a construct is intended to include these, for present purposes. Insofar as any of these operations reflect the

play of some variable established as such through a specifiable history of conditioning, it appears appropriate to refer to the application of that variable as a conditioned drive operation, if only to maintain a certain consonance with traditional usages of the term drive.

This concludes a brief outline of the conceptual scheme, and the hypotheses it generates, which is to be applied to the problem of the dimensionality and the development of emotional dependence in childhood. Finally, it should be noted that this system is intended to be open to modification and addition; where new constructs prove necessary to account for variances not accounted for by the basic constructs employed, in principle they *must* be developed.

Pertinent Theoretical Statements

The development of emotional dependence has been the subject of relatively little formal theorizing. Speculation has sometimes taken the form of extending to some aspects of dependence existing theoretical schemes which had proven useful in ordering relatively simpler learning phenomena (e.g., 1, 8), as is the case in this paper, and sometimes the form of describing and interpreting clinical case histories in terms of emotional dependence. While considerations stemming from many sources have been taken into account in the preparation of this paper, limitations of space do not permit their detailed examination.

The Basic Plan of Research

The overall research program is to proceed essentially from the simple learning and performance concepts set forth in earlier sections toward tracing the development of emotional dependence in children. The most fruitful application of these concepts demands that certain priorities be established among the many classes of research which they can generate. Initially, a two-pronged approach appears suitable: the *dimensional* program can begin with an examination of the characteristics of behaviors typically labeled emotional dependence; while the *developmental* program can focus upon their antecedents, both at simple and complex levels of analysis.

The Dimensional Program

A possible strategy in the research on early dependence learning would be first to establish a catalogue of the characteristics of emotional dependence, the details of which might then be employed in the researches of the developmental program to focus effectively upon the events which typically operate in that early learning. This program would begin with an enumeration of the potentially limited number of social reinforcers, of responses employed for those reinforcers, and of object persons who can provide the reinforcers, all of which characterize the process in young children. Next studied would

be the relationships holding among responses, reinforcers, and objects, and the distribution of hierarchies of efficacy of the reinforcers for various populations. Given the beginnings of this catalogue, this program finally would add developmental level, defined in terms of age, as one of the basic parametric variables with which it will deal, thereby ordering emotional dependence as a developmental process. While reflecting the cumulative history of the organism, age provides a crude index of a number of developmental processes, and represents a dimension correlated to stages in development at which qualitative changes occur. The mechanics of these variables are dealt with in the program which is described in a later section.

Some preliminary work relevant to the dimensional problem can illustrate some of the above points. In an investigation of some parametric and dimensional aspects of the attention-seeking process, it was found possible to distinguish nine responses which were assumed to be employed for an adult's attention, and some perhaps also for his approval (2). These could be reliably scored, and entered into a pattern of relationship with the independent variables. Further, it was found that the six responses which were verbal in character clearly defined a factor in a factor analysis (3); and that when the distribution of frequency scores for each response was dichotomized about zero frequency, and the two responses with the most zero scores were combined, a five-item Guttman scale was produced (4). The responses composing the scale, according to their increasing scale values, are: Comments (casually emitted), Questions (casually asked), Attention-seeking (urgent responses for overt notice from the adult), Help-seeking, and Permission- *or* Praise-seeking. Thus, it appears that a dimension of behaviors apparently employed for the reinforcer of attention, and perhaps also for approval, has been identified, and can serve as a first approximation for ordering dependence responses. In addition, age yielded almost no reliable relationships with the nine responses in the four- to six-year age range employed. However, when Heathers (5) compared the free play behaviors of 20 two-year-olds and 20 four- and five-year-olds, he found that clinging and affection-seeking decline with age relative to attention- and approval-seeking, and that emotional dependence on adults (nursery teachers) declines with age relative to emotional dependence on other children. The writer has found that children age four to six exhibit relatively greater frequencies of responses for attention to adults of the opposite sex as themselves than to adults of their own sex, with the difference much more pronounced in boys than in girls (2).

These results indicate that the goal of isolating and cataloguing a variety of responses, reinforcers, and objects involved in dependence behaviors, and of stating this catalogue as a function of developmental level, is reasonable.

Some further experiments were conducted under the writer's auspices to evaluate the effects of reinforcers which, in the setting of the experiment, presumably appealed to the reinforcer of approval. They were carried out

also to provide, in the differential susceptibility of *Ss* to such reinforcers, a dependent variable which might yield considerable leverage upon the characteristics of emotional dependence. The reinforcers were found to exhibit a surprising complexity: for while it was established that a number of simple responses in children of various ages (e.g., bar-pressing, marble-rolling, letter-naming, word-emission) could be controlled effectively by the casual use by the experimenter of terms like "good" and "uh huh" (which control was in the main unrecognized by the *Ss*), it was also found that apparentaly identical control (also unrecognized) was achieved by using for a reinforcer a sound (that of a penny dropping into a bottle). Unlabeled learning or learning without awareness may provide a prototype for early identification learning in children.

The Developmental Program

The developmental program describes researches designed to supplement the dimensional program by establishing the laws underlying the development and subsequent efficacy of the generalized reinforcers, and possibly the conditioned drives, assumed to be involved in emotional dependence. In effect, this requires the eduction of some specific laws of learning to the special case of the developing human, and the initial focus upon the earliest dependence learning susceptible to experimental investigation. In order to accomplish this, a two-pronged developmental research program designed to converge more or less simultaneously upon different levels of problems is outlined: The first will investigate the parameters of generalized reinforcers, employing relatively simple (i.e., stripped down) learning situations to elucidate the processes involved; and the second will deal (perhaps at several levels of analysis simultaneously) with the actual details of the presumably complex learning of emotional dependence in the young child. While the latter is emphasized, clearly the two subprograms are interdependent.

The Basic Characteristics of Conditioned and Generalized Reinforcers. Researches of the first subprogram would identify the variables which determine the strength of conditioned and generalized reinforcers, and then study those variables as parameters. Thus, there would be studied the effects of such training variables as: number of pairings of conditioned with primary reinforcers; schedule of such pairings; temporal relation between conditioned and primary reinforcer; magnitude of primary reinforcer; and number and type of primary reinforcers, and of deprivation states involved; and different effects on reinforcers conditioned by positive and negative primary reinforcers. In addition, selected performance variables (e.g., different deprivation conditions) under which conditioned reinforcers are operative, as well as the effects of the deprivation of conditioned reinforcers, would be studied. The basic information provided by such researches could serve as one class of hypotheses for the study of the details of the child's

emotional dependence learning. Fortunately, research on basic learning continues to produce studies of this class.

Generalized Reinforcers and Emotional Dependence. Departing from the simplest level of analysis that the first subprogram represents, one can move as freely among levels of complexity as is required by the problem at hand. Stimulus events more social than, e.g., the lights or buzzers employed in studies of the first subprogram may be investigated as discriminative stimuli and conditioned reinforcers (e.g., physical appearance, attention, or approval of a caretaker), as may be relatively social responses (e.g., "Look at me!"). In addition, the variance in the potency of a number of presumed generalized reinforcers already effective for children in the early years, can be studied in relation to specific primary deprivations, some of which were involved presumably in early dependence learning. Other deprivations can be of the generalized reinforcers themselves, to determine whether the effects produced are similar to those produced by primary reinforcer deprivations, or by other drive operations. If such operations were to meet the criteria, they would be termed conditioned drives. The results of such researches could be of considerable utility, for they might point to at least two qualitatively different types of early learning, one producing generalized reinforcers alone, and the other producing conditioned drives as well. If this possibility eventuated, the identities of the events falling into these two classes, as well as the identities of the primary deprivations to which they might be sensitive, would yield valuable clues about early childhood learning.

At times, loosely-defined dimensions such as "rejecting-accepting," "warm and nurturant—cold and neglectful" and the like have served as antecedents of a variety of social behaviors in theories and descriptive laws; they have described long term patterns of behavior of caretakers towards their children. It could prove useful to the program outlined, and perhaps serve to articulate it with a number of extant theories and researches, if some of these gross concepts were examined, and the reinforcement histories involved (e.g., schedule of reinforcement, identity of reinforcers) postulated.

As a unit of the dimensional program, an attempt to determine the distribution of hierarchies of reinforcer efficacy is to be made, and groups of *S*s found homogeneous in this respect established. Given such information, it would be possible to attempt to establish the identities of such groups according to criteria as, for example, might be suggested by the conceptual scheme outlined, or by others. Communalities in antecedents could be identified in this manner, and information valuable for the developmental research program obtained.

Concurrently, researches on an infrahuman level may be conducted to examine the effects of postulated variables for the special case of human infancy. In particular, research on this level could be useful in studying

variables too extreme or too vaguely understood to be well implemented with humans, as well as for certain variables involved in long term longitudinal studies.

A Possible Experimental Program with Infants. Finally, a variety of researches on human infants under such caretaking conditions as would be provided by a highly controlled institutional or home setting would become feasible and desirable as data accrued from the dimensional and developmental researches. These would be either in the form of experiments which would attempt to produce particular effects subsequently in the child's pattern of emotional dependence (reinforcer and response hierarchies) by varying selected aspects of the caretaking process; or they might employ systematic observational techniques to assess the learning contingencies for the child, which could be referred subsequently to his pattern of emotional dependence.

The experimentation would be aimed at producing a prototypic "model dependence" in young children. Both positive and negative reinforcing aspects of caretaking may be employed as variables. Selected stimulus events involved in or attached to the caretaking person may be set into a variety of contingencies with these aspects of the caretaking process; and selected aspects of this process may be made contingent upon different responses emitted by the child.

By postulate, such procedures will imply consequences in the later behavior of the children to whom they are applied, and the experimenter may find that he has control of the laws governing some major dimensions of social interaction. Researches initially would be restricted to the use of contingencies holding between socially trivial stimuli such as lights or buzzers, and socially trivial responses such as head rolling, rocking, or playing with a toy. The range of control of the caretaking process must be established cautiously. In instances where it is difficult to establish clearly the nature of the variables to be applied, or the most efficient method of applying them, the problem may be referred to analogous experiments on animals. Responses such as head rolling or rocking, aimed at securing "model" reinforcers such as lights or buzzers, either from people of from a simple mechanical apparatus capable of their delivery, would be studied. Hence it may be possible to speak of "light-seeking" as a "social" behavior or simply as a general behavior. The effects of other parameters, such as the scheduling of the contingencies, the differential aspects of stimuli, caretaking, and caretakers involved in those contingencies, and such general parameters as age, or pre- and postverbal learning, may be applied both for the general case and for the "social" case.

These relationships need not be restricted to the "model" level described above, for neither the stimuli nor the responses involved need be trivial. Such events as the attention, approval, or affection of the caretaker may be the principal stimuli involved, and responses of social significance, viz.,

attention-, approval-, and affection-seeking, may result. Validation of the model described must be affected cautiously, and it may be possible to conduct such researches only in a late stage of the program.

Some General Problems

It is expected that many of the behaviors of young children may vary with developmental level. Any research program aimed at developmental functions would do well to anticipate such of these as appear relevant. For example:

1. The hierarchy of reinforcers for the young child may vary with developmental level. A portion of this variance may be attributed finally to variables investigated in the programs described here, while another portion of the variance may result from very general variables at present not of immediate relevance to those programs, e.g., the types of discipline employed by the parent at various ages. Such effects may be reflected in the proposed catalogue of reinforcers, and could be evaluated there as such.

2. The efficacy of a given set of learning procedures may vary with developmental level. For example, it appears likely that preverbal and postverbal learning differ in several qualitative aspects. Thus, it is possible that two individuals to whom apparently identical conditioning procedures have been applied will show different sensitivities to the reinforcers involved, simply because the procedures were instituted at different developmental levels.

Other problems or issues of a general nature which appear relevant to the program are:

3. It is possible that certain inevitabilities of human caretaking will produce some response-reinforcer correlations of a very general occurrence. For example, it may be discovered that attention is an inevitable reinforcer for the response of talking. Hence, it may be difficult to account for such correlations initially, since there may be relatively little variance apparent in the population, and experimental control may be impractical. Consequently, the catalogue of reinforcers may appear to contain "unexplicable" relationships.

4. The acquisition mechanics utilized in conceptualizing this problem allows a wide range of possible stimulus events to meet the criteria defining the development of generalized reinforcers pertinent to dependence behaviors. However, certain aspects of the typical child-rearing situation yield a number of expectations about the identities and interrelationships of these stimulus events, which may operate to reduce substantially the number and range of the significant generalized reinforcers typically conditioned. Hence, these expectations must be examined and evaluated for their implications.

For example, it is thought that before a child is reinforced by his caretaker, certain stimulus characteristics of that caretaker typically will be made

apparent. In the general case, it may be anticipated that caretaking is *most* probable from a caretaker who is at once (*a*) close at hand, (*b*) attentive to the child's state of need, (*c*) approving of his request for care, and (*d*) affectionate in the delivery of that care; on the other hand, caretaking is least probable from a distant, inattentive, disapproving, disliking person. Hence, there can be postulated a continuum of the probability of caretaking as a function of the differential presence of these stimuli, which (according to the acquisition mechanics utilized) implies consequent sets of hierarchies of these same stimulus events as generalized reinforcers controlling dependence responses. The following appear to be some major hierarchies which may result from typical caretaking conditions:

i. The proximity and attention of the caretaker may be the stimuli most strongly correlated with the caretaking process. This is probable particularly in the case of the very busy caretaker, who can be near the child only to see to his needs, and must be elsewhere at other times. However, in an environment of very limited space, the proximity of the caretaker may be inevitable and continual. If, in this situation, the caretaker is very busy, then her attention may be the single stimulus most strongly correlated with caretaking.

ii. The proximity and attention of the caretaker will not be so strongly correlated with caretaking in a situation where she has greater leisure, for then her presence and attention will be available on many occasions which need not be caretaking occasions. Thus, it is likely that there will be times when the child may receive care of a minor nature. In such situations, it is anticipated that other characteristics of the caretaker, e.g., her approval or affection, may fulfill the criteria of discriminative stimuli more adequately than will her proximity or attention.

The Methods of Research

The research program that has been proposed is both diverse and intensive. Hence, ideally it would require highly detailed investigations of a considerable number of problems. It is not appropriate here to detail experimental designs for these researches: to some extent they are a function of the data accruing early in the program, and hence must await those data. Consequently, it is practicable to outline only the basic methods which will probably be required in the overall program.

It is recalled that an initial problem is to recognize and identify the social reinforcers which characterize emotional dependence behavior. The definition of a reinforcer provides an immediate criterion for this problem: it is a stimulus event which, if made contingent upon some response, yields control of the increasing strength of that response. However, this definition should be entertained only simultaneously with the general expectation, noted earlier, that some reinforcers may be pertinent more to certain

responses than to others. In general, this initial problem of defining the responses and reinforcers comprising the class of emotional dependence behaviors requires some type of standard situation, within which it is possible for *S*s to emit responses upon which reinforcers may be made contingent. This approach, with suitable variation, allows for the definition of social reinforcers, and for some assessment of their possible correlations and hierarchies.

The Baseline Method. The introduction of this method into psychology is due primarily to the work of Crozier and Skinner. Its logic is simple, and may be applied efficiently to experiments required in the program outlined. In terms of statistical design, the method essentially represents a treatment by subjects design (perhaps including a variable representing successive units of time or trials), in which a highly dependable response variable is employed.

In principle, the logic of the method is as follows: if the rate of output of any instrumental response can be elicited dependably under standard conditions, those data would provide a baseline against which the effects of treatments could be evaluated efficiently. When a stable rate of output is attained, for the purpose of an experiment a major portion of output variance, perhaps with the exception of variance due to differences between *S*s, is under experimental control. Thus, a highly efficient experimental method is provided, one which generally requires small numbers of *S*s. Considerable exploratory work typically is required to establish conditions under which the highly reliable dependent variable is generated.

The nature of the response selected would depend on the given problem. A simple response may be employed, even in a program whose focus is the study of complex human processes. The experimenter's (*E*'s) logic might content him with the laws involving the simple response, or he might choose to validate some of those laws against laws which involve the more complex behavior; and then, if a high degree of correspondence were found to obtain, he could continue to employ the simple response as his most efficient vehicle for gaining leverage on the complex behavior class. Alternatively, he might employ the complex response from the beginning.

The steps in the typical experiment are very much as follows: First, a response (e.g., bar-pressing, naming-things, attention-seeking, etc.) is selected. Then, a dependable rate of output is established under a given set of conditions, such as that generated by a particular schedule of reinforcement, or set of instructions. If the dependent variable so generated can take a wide range of values, it can provide the required baseline. After the constant rate of response output has been developed in *S*, a treatment is introduced, and the response rate observed. When the rate has become stable under the treatment, the reversibility of the process can be examined, as can the efforts of different degrees of the experimental variable, allowing perhaps for the examination of the complete range of the variable in one *S*.

The number of repetitions necessary to establish its reliability would depend on the magnitude of the effect and the consistency of the change in rate of response from treatment to treatment. The experiment is repeated on a number of *S*s sufficient to establish that the effects of the treatment are significantly large with respect to individual variability. Few *S*s are required if the output records indicate that the response effects vary in the same way from *S* to *S* (and differ perhaps only in terms of a constant effect, e.g., a *S*'s level of responding).

Some Illustrative Techniques. To date, employing the baseline logic, there has been developed to a nearly satisfactory stage a method for defining social reinforcers, when behaviors to attain them are channeled into one of two motor response outlets. The child (*S*) is introduced to the setting as to a game, the central aspect of which is a toy displaying two holes down which marbles may be dropped (to return at the base of the toy). *S* is given a supply of marbles, shown briefly how to play the game, and allowed to establish baselines for the use of each hole. Subsequently, *E* chooses the nonpreferred hole, and makes some social reinforcer contingent upon its use. This contingency may be announced to *S* or not, depending upon the formal characteristics of the reinforcer utilized. Consequent shift from baseline use of each hole indicates the degree of control that the reinforcer involved can exert, for a given *S* on a given occasion. In principle, it is possible to articulate a variety of closely discriminated social reinforcers into this situation by tailoring the set appropriately.

For example, *S* may be introduced to the toy, and to the recording apparatus housed in an adjoining room. He may be told that the toy runs the recorder. After his baselines have been established, *E* may announce that he can stay and watch *S* play only so long as marbles are rolled down a certain hole; otherwise, he must see to tasks (e.g., tend the recorder) which requires his attention in the next room. Degree of use of that hole by *S* in excess of his baseline use of it assesses the potency of the social reinforcer involved in this set, which tentatively may be labeled as *attention involving proximity*. An alternative set, which appears to appeal primarily to *approval*, is constituted by *E* remaining in the room with *S* and, subsequent to the establishment of the baselines, pointing to one of the holes and remarking, "You know, this a the good hole." This set can be elaborated by *E* expressing approval subsequently whenever the hole is used (e.g., by saying "Good!," "Fine!," "Uh huh," or "That one was good"). An alternative set, appealing in its face characteristics to *achievement* as well as approval, can be constructed from the approval set by *E*'s counting aloud the number of "good" responses which *S* makes. Alternatively, achievement may be suggested more clearly by telling *S* that in this game he can get "points" if he wants, and then dispensing tokens, or making a tally mark whenever *S* has employed the correct response a sufficient number of times. Here the game element can be preserved by letting *S* discover the identity

of the hole which yields points. Again, the ability to *internalize* approval or achievement may be assessed by *E* leaving the room after the approval set has been established, thus leaving *S* to maintain the correct response (if he will) alone. Processes such as *identification* may be assessed by establishing a set in which *E*, after noting *S*'s baseline, himself takes a turn at playing the game (as a friendly demonstration), but restricts himself almost completely to the use of one hole. Consequent copying behavior on the part of *S* may qualify under a given experimental logic for the label of identification.

From these examples, it is clear that a large number of simple, natural sets can be constructed which have a face appeal to the variety of stimulus components typically thought to constitute the reinforcers involved in emotional dependence. Experience suggests that each *S* may be used again and again, under the same or different sets, if these occasions are separated in time by a few days. Hence in principle it is possible to establish efficiently and in fine detail many of the intercorrelations, equivalencies, and hierarchies of these social reinforcers. Since it has been found, to date, that children as young as four perform satisfactorily in such settings, these interrelationships may themselves be stated as function of a developmental parameter indexed by age. Such findings would have considerable relevance to a number of theoretical postulates about the development of these behaviors; for example, need a child be strongly responsive to approval before he will display behaviors labeled as identification, or achievement?

By the same token, a wide class of nonhistorical antecedent variables may be articulated into these experimental sessions, thereby establishing laws which may be basic to everyday social interactions. This possibility is quite reasonable when it is recalled that, while the experimental situation involved here is a skeletal, "stripped down" setting, nevertheless the stimulus events delivered as social reinforcers represent a fair sampling of the reinforcers utilized in life settings; similarly, the antecedent variables applied also may be drawn from a class of typical life events. For example, an *S* may be run under any set (i.e., for any social reinforcer) immediately after a period of social isolation. This variable can be affected in practical terms by isolating *S* in a room, with a toy to play with or a task to solve, and leaving him there for as long as is practicable. Any effects which this procedure will reflect in the typical potency of the reinforcer subsequently dispensed in the experimental session can be of great relevance in accounting for variance in behaviors for such a reinforcer in everyday life: whether a period of social isolation is conceived of as a condition of deprivation or as a change in the schedule of receipt of social reinforcers, still it parallels the fluctuations in the availablity of social reinforcers typical to an everyday social environment.

In a similar manner, many classes of social variables could be abstracted from complex environmental contexts and applied to these experimental

sets, and their effects assessed in the consequent responsiveness of responses for the social reinforcers dispensed there. As examples, consider sex of the reinforcer-supplying adult relative to the sex of the child, selected primary deprivations, and the definable status groups (peers, etc.) from which spectators are drawn.

However, it must be remembered that in a less structured situation where the child can choose his "reinforcer-seeking" responses, it may well be expected that some responses may be more pertinent to some reinforcers than to others. To assess these possible reinforcer-response correlations, a standard situation like that employed in Reference 2, leaving the child free to show a wide range of responses, is required. (The above situation which channels all responses into a single motor response clearly is not applicable to this problem.) It is also possible to institute in such a situation, e.g., two or more levels of the availablity or supply of such reinforcers, and hence the interrelationships of the dependence responses emitted can be reflected against such a basic variable. Such comparison allows deductions about the essential nature of many of the reinforcers, e.g., their equivalencies or hierarchies under conditions of relative commitment of the adult as a source of reinforcers.

Other Procedures

The results of factor analyses may serve to demonstrate independent factors operative in the phenomena of emotional dependence, which possibly can be referred to distinct prior processes. Similarly, the use of some recently developed scaling methods for qualitative data, and in particular Lazarsfeld's method of latent structure analysis (7), may serve to define groups of *Ss* which are homogeneous with respect to different hierarchies of generalized reinforcers; and if it were possible to establish the identity of these groups in terms of such criteria as distinct prior learning contingencies, then a number of new research avenues would be opened. The range of independent criteria on which such groups might differ is broad, and hence is best left to the reader's imagination.

Summary

A research program is outlined for investigating the laws characterizing emotional dependence in childhood. In particular, emphasis is placed upon the identification of the dimensions operating in behaviors which appear to be employed to gain certain social responses from other persons (like those apparently employed for attention, approval, affection, reassurance, and the like); and the conditions under which they develop. The program proceeds both from simple learning and performance concepts and from such theories as are available for tracing the early development of emotional dependence.

Initially, a two-pronged research program is proposed: the *dimensional* program would catalogue instrumental behaviors, reinforcers and object persons which characterize the dependence process, and such interrelationships as may be discovered among them, stated as a function of age; and the *developmental* program would employ the catalogue to focus upon the etiologies of different emotional dependence patterns. In particular, these researches propose to focus systematically upon selected aspects of the human caretaking situation.

References

1. Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
2. Gewirtz, J. L. Three determinants of attention-seeking in young children. *Monogr. Soc. Res. Child Develpm.*, 1954, 19, No. 2.
3. Gewirtz, J. L. A factor analysis of some attention-seeking behaviors of young children. *Child Develpm.*, 1956, 27, 17-36.
4. Gewirtz, J. L. Steps toward a scale for children's attention- and approval-seeking. In preparation.
5. Heathers, G. Emotional dependence and independence in nursery school play. *J. genet. Psychol.*, 1955, 87, 37-57.
6. Keller, F. S., & Schoenfeld, W. N. *Principles of psychology*. New York: Appleton-Century-Crofts, 1950.
7. Lazarsfeld, P. F. Latent structure analysis. In S. A. Stouffer *et al.*, *Measurement and prediction*. Princeton, N.J.: Princeton Univer. Press, 1950. Pp. 362-472.
8. Skinner, B. F. *Science and human behavior*. New York: Macmillan, 1953.

SOME RESEARCH PROBLEMS FOR DEVELOPMENTAL PSYCHOLOGY

EUGENE S. GOLLIN
Institute of Child Welfare
University of Minnesota

The function of this paper, as outlined by the organizers of the conference, is to present matters of high research priority in the area of developmental psychology. Such presentation is bound to reflect theoretical, methodological, and personal biases; it follows that many matters of high priority are thus excluded. It is with this recognition that the following considerations are submitted.

Reviewers of the field of child psychology have been relatively pessimistic of late with regard to its general status. General pessimism was expressed by Barker in the second volume of the *Annual Review*. He stated that the field was lacking in vigor (2). By the time the fifth *Annual Review* was prepared the problem was no longer one of the vigor of the field but rather was concerned with the lack of rigorous methodology and systematic theory (17). Other criticisms have been directed toward the apparent lack of integration between child psychology and general psychology, and the use of the child as a special subject for study in his own right (23).

The former concern operates both ways. For example, while some feel that child psychology has ignored general psychology, others assert that general psychology has been slow to take advantage of developmental principles (44).

The utilization of the child as a subject for study in his own right has applied value, but researchers involved in the pursuit of general principles have the responsibility of viewing their developmental samples within the encompassing framework of phylogenesis and ontogenesis. It ought to be

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

recognized that child psychologists are not the only psychologists who from time to time abandon this perspective. See for example Harlow's caustic comments on "rodentological learning theory" (10), Beach's remarks about the narrowness of species sampling in general experimental psychology (3), and Verplanck's call to observe more species and observe them under a wider variety of behavioral conditions (including ontogenetic periods) than those generally utilized in the laboratory (40). These criticisms are directed at general and experimental psychology and would indicate that much of the pessimism expressed about child psychology should be generalized to those fields.

The outline of a high priority research program is inherent in the dissatisfactions cited above. The theme of such research must be *change* in the sense that Sears has employed the term. Sears discusses two aspects of behavioral change: one involves ongoing action, or performance, and the other consists of learning, or acquisition (36, p. 478). Ongoing action reflects those properties of a person that describe his potentialities for action. Among these properties, according to Sears, are drives, habits, cognitive structures, and expectancies. The study of personality development requires a set of principles that will describe the way in which these properties, the potentialities for action, are changed by experiences. The changes in potentialities for action are best described by the laws of learning.

If Sears' concept of change is elaborated so that it explicitly includes behavioral manifestations that indicate shifts in performance and acquisition which are regressive, deteriorative, and declining in rate and proficiency, then it will prove to be a suitable frame-of-reference for the kind of developmental psychology called for by Dennis; a psychology adequate to the course of human life as a whole (4). The investigation of process, within this framework, includes utilization of geriatric and pathological samples, and the employment of laboratory techniques designed to impair as well as enhance performance.

A comparative investigation of ongoing action and acquisition raises questions not only concerned with the way in which the latter alters the former, but also is concerned with the influence exerted by the properties of ongoing action on the course of acquisition. A comparative approach necessarily emphasizes organismic characteristics, and centers interest on differences as well as similarities in the basic mechanisms of modifiability. The nature of the organism assumes prime importance in comparative investigation. That this is not always the case is amply demonstrated in Postman's operational definition of a perceptual experiment. The operations which define the perceptual experiment include (1) manipulation of the stimulus along selected physical dimensions, (2) certain specific manipulations of the organism and, (3) the measurement of a limited class of response (31). Comparative approaches add to this definition certain specifications, crude or otherwise, concerning the nature of the organism that

provides the measured responses. Such definition is sometimes attempted by describing antecedent events. It is possible that description of this sort is ultimately adequate. In practice, however, it is likely to be quite incomplete; generally it is confined to laboratory manipulations, in which case it becomes synomymous with (1) and (2) above. Hilgard recognizes this problem directly when he asserts that paranoids do not use their past experiences the way other people do when they interpret what is going on in the present. He points out that gestaltlike principles may be necessary in accounting for how contemporary organization draws upon past experience in the achievement of perception (13). A similar argument is presented by Luchins who questions the assertion that all organization is based on past experience and sees the need for experimental inquiry into the principles governing the acquisition and operation of experience (22).

The potency of frequency, frequency of reinforcement, contiguity and like parameters for behavior is not denied when organizational factors are introduced. The function of such introduction is to direct experimental attention to certain divergencies between cognition and neo-association theories. A central problem posed by these divergent viewpoints involves the place, if any, in behavior theory, of what cognition theorists have called central representational processes. The adequacy of the classical Gestalt rendtion of these central representational systems or schema has been vigorously questioned from a number of directions. Piaget questions the Gestalt assertion of the permanence of certain forms of organization throughout mental development (29). Hebb discusses innately given, sensorially determined unities in perception and supplements these with nonsensory unities and identities primarily dependent upon experience (11). Harlow introduces experimental evidence which raises questions about the "equilibrium" assertions of Gestalt theory (10). Krech critically analyzes the sufficiency of the law of pregnanz for the description of behavioral events (19). This is but a sampling of the broad theoretical attack being levelled at orthodox Gestalt theory. This attack creates a number of problems for research. The attackers by no means rule out such things as figure-ground organization, progressive differentiation, or the notion of levels of functioning. Hebb retains *figure-ground* as an important conceptual tool, and in later writings (12) deals with motivation in a manner reminiscent of Lewin's discussion of valence (20). Krech amends the law of pregnanz to include the spontaneous development of states of greater heterogeneity and complexity, but his *Dynamic System* remains a field complex (19). Piaget retains and expands his position on the importance of schemata as determinants of behavior. In discussing the nature of the assimilatory act, he stresses the importance of differences in behavior patterns which characterize different periods of development (29, pp. 150-153). Assimilation is defined as the action of the organism on objects which surround it (29, pp. 4-8). There is a progressive differentiation of operations. Early operations, genetically speaking, are bound to

immediacy, to that which is directly given in perception. More lately emerging operations permit conceptualizations which involve accommodation to contexts not directly given in perception (1, p. 30; 30, p. 189 ff).

Werner's system represents a similar modification of classic Gestalt theory. He describes development in terms of a series of organizational levels (43). Primordial behavior tends to be syncretic and diffuse. It is succeeded by gradually increasing differentiation and hierarchic integration. At both early and mature periods of development organization tends to be wholistic (5, 44). Primitive wholism is characteristically undifferentiated whereas the more mature whole manifests good articulation of parts and integration of parts into an encompassing system.

Murphy's discussions of development are akin to those of Piaget and Werner. Murphy stresses the sequential appearance of global, differentiating, and integrating function patterns, and points out that these patterns manifest themselves in growth, and in learning in both genetic and agenetic situations (27, pp. 66-68; 28, pp. 342-343).

Scheerer, also, suggests that representational modes active at various periods in development may retain their functional potency and reassert their influence under certain special sets of conditions in the adult (35, p. 9).

A conception not dissimilar in form to that offered by the cognitive writers is Hebb's description of the development of perceptual learning. He states:

> The course of perceptual learning in man is gradual, proceeding from a dominance of color, through a period of separate attention to each part of a figure, to a gradually arrived at identification of the whole as a whole: an apparently simultaneous instead of a serial apprehension (11, p. 32).

The central representational processes or schemata may be thought of in terms of Sears' "ongoing action or performance." For practical experimental purposes they must be considered as limiting factors in behavior. The refinement of certain orders of performance should, if such schemata are operative, be restricted along predictable lines. Systematic training series, Sears' "learning, or acquisition" variables, should produce differential learning effects in populations defined in terms of inferred organizational characteristics. Such population definition may proceed in terms of predetermined criteria, e.g., chronological age, intelligence, conditions of pathology, or in terms of performance in special test procedures utilizing tasks specifically designed to produce performance variations. Training procedures, for heuristic purposes, ought to be conducted along the lines suggested by systematic learning theories.

Tests employed to differentiate or categorize individuals with respect to organizational characteristics must be adequate to the task. Razran draws a distinction between "mere acquisition" of conditioned reactions in which perception is unnecessary, and acquisition under conditions involving perceptual relations. He states:

> Or in other words, whenever learning takes place, it involves the operation not of one kind *or* the other kind, but of either (*a*) the lower kind or (*b*) the higher *and* lower kinds. And specifically it obviously means that relational learning subsumes conditioning. . . . (32, p. 92).

Razran's discussion is pertinent to the question of the selection of test materials because the schemata or representational systems proposed by cognitive theorists are testable only when the task fulfills condition (*b*) above. Learning situations designed to subject the inferred system to test must involve the operation of relational factors. This point was clearly made some time ago by McGinnis in her study of the motor habits of young children (24, pp. 289-290). She found no relationship between first trial performance in a pattern maze and first trial preformance on a concealed-cue maze. While first trial performance on the pattern maze was related to intelligence, no such relationship was found between first trial performance on the concealed-cue maze and intelligence. McGinnis' pattern maze (the Young Slot Maze) is adequate for testing inferred central representational systems since first trial learning involves the discernment and utilization of path relationships. The concealed-cue maze is not an adequate testing device since first trial learning involves strictly rote procedures.

As an initial step in setting up a research program directed at exploring the relevance of cognitive theory for behavior, a series of base performance characteristics must be obtained on variously defined population samples. This is particularly necessary since cognitive theorists have tended to be relatively vague concerning a number of their constructs. A programmatic research example of the specification of organizational properties is provided by Werner (44), Friedman (5), and Siegel (38). Utilizing the Rorschach blots, and concentrating on structural features, a genetic scoring system has been developed. For example, two kinds of *W* responses are distinguished:

> . . . genetically low responses are those which indicate amorphous, diffuse, or confabulatory percepts where little attention is given to part relations and to perception of contours. The genetically high percepts are reflected in the responses whereby the percept is that of a precisely formed unit with parts integrated into it, where the whole is composed of relatively independent sub-wholes. . . . with age there is a decrease of the undifferentiated diffuse whole and detailed responses, there is an increase of highly articulated, well-integrated whole and detail responses. There is further interesting shift from the early whole responses toward small details between the ages of about 6 to 8 years declining in favor of the integrated whole response later on (44, p. 4).

Friedman reports that schizophrenics exhibit features of structural regression in the Rorschach which resemble the perceptual performance of normal children (5). Siegel, working with subjects characterized by degree of impairment, i.e., paranoid, and hebephrenic and catatonic schizophrenics, reports that the paranoid group produces Rorschach responses similar in

structure to those obtained from normal children between the ages of six and ten years, while hebephrenic and catatonic subjects produce responses similar in structure to those found among normal children in the three to five year age group (38). Werner reports a series of "microgenetic" changes which resemble the ontogentic changes described above. When normal adults are presented with Rorschach blots under exposure times of 1/100 of a second, 1/10 of a second, 1 second, and 10 seconds respectively, the structures of the perceptual responses parallel those obtained from ontogenetic samples under conditions of prolonged exposure.

These Rorschach studies provide not only increased specification of cognitive inferences, but also suggest a methodology for integrating developmental, general, and clinical psychological problems within an experimental framework. This framework will prove fruitful if it is expanded to include parameters usually emphasized by systematic learning theories.

An experimental program along these lines is currently being launched at the Institute of Child Welfare at Minnesota.[1] It is concerned primarily with the development of visual and haptic recognition. Recognition occurs when an appropriate designating response is made by a subject to a stimulus with which he has had previous commerce. Such designating responses are not confined to exact stimulus reproductions, as shown by the extent of recognition behavior in everyday life. Individuals recognize many objects with which they have not had direct experience. Persons also frequently recognize objects which are partially obscured, disguised or altered with respect to their original representation. E. J. Gibson's review of pertinent literature presents investigations which cover impoverishment of stimuli, effects of practice and reinforcement, and transfer. She concludes that there is ample evidence that perceptual learning occurs under many conditions (6).

Despite the vast array of data on conditions under which discrimination and recognition are enhanced, we know little concerning the development of these behaviors. What are the governing principles? Association theorists propose that behavior in general, including recognition, develops gradually and is subserved by reinforcement (14), contiguity (9), redundancy (26), etc. Cognition theorists introduce the notion of saltation in development and assert that behavior occurs as a function of the underlying patterning of psychological processes. Differentiation of psychological function imposes limits on the nature of performance, and behavior is accounted for in terms of the inferred degree of differentiation (18, 20). More extensive descriptions of developmental changes in central patterning are provided by Murphy (27), Piaget (29, 30), and Werner (43). The assertions contained in these alternate points of view suggest an experimental series in the area of visual and haptic (stereognostic) recognition which may serve to clarify particular theoretical issues.

[1] This research is aided by a grant from the National Science Foundation.

Initially it is important to select experimental tasks that contain selected features and are capable of eliciting wide arrays of individual performance differences. This procedure is accompanied by the selection of subjects in terms of (cognitive) theoretically-inferred characteristics. Characteristics of developmental samples have been specified both by cognition theorists, and others (25). The selected test materials should elicit performance differences consistent with the predictions made about the several subject groups. Once such base performance rates are available specific tests can be made concerning the behavioral effects of the parameters of systematic learning theories. Subjects are drawn from virtually the entire age range, represent various levels of intelligence, and degrees and conditions of pathology. This basic procedure is to be replicated with a number of tasks and in several perceptual modalities.

A pilot study exemplifying the initial phase of basic procedure has been completed. Subjects in this study include 48 preschoolers from the Nursery School and Kindergarten of the Institute of Child Welfare and 17 adults drawn from graduate courses. The materials consist of 15 common objects, e.g., umbrella, automobile, ice cream cone, etc. These objects are presented on cards, one at a time, and identification is requested. The contours of the pictures are produced by a series of small cross-hatches and obscured or embedded by crosshatching over the face of the card. After familiarization trials, *S*s of preschool age received exposures of four seconds duration while for adults the exposure time was one second. Table 1 presents the mean number of correct identifications for the several age groups.

TABLE 1

MEAN NUMBER OF CORRECT IDENTIFICATIONS OF MEANINGFUL EMBEDDED FIGURES BY AGE GROUPS*

Age	*N*	*Mean*
2-6 — 3-5	17	1.00
3-6 — 4-5	18	2.78
4-6 — 5-5	13	5.23
Adults	17	11.53

* Ages 2-6 to 5-5 received four-second exposures while adults received exposures of one second.

The beginnings of a base performance rate for the particularly characterized groups begins to emerge. Alternate sets of explanations for these performance differences are available from association and cognition theories. The latter place emphasis on differential figure-ground facility while the former stress such aspects as frequency of prior exposure, or reinforce-

ment, etc. Systematic training series with these materials is currently being undertaken. The effects of training will reflect directly on the various theoretical assertions which have been reviewed.

Another experiment, utilizing different materials, has provided similar ranges of performance variation among preschool populations (7). Once again common objects were used as stimuli. Each object was represented five times to constitute a series involving a gradation of completeness. The initial card contains very few contour lines or internal details and these are gradually built up through the series until on the final card the complete object is present. There are eleven such series of five cards each. Only *Ss* who have the appropriate verbal designating response for the complete pictures are included in the analysis. Even under these conditions there are marked differences among age groups in terms of the degree of completeness of representation required before recognition occurs.

A similar line of investigation is being carried forward with geometric figures under conditions of embeddedness. Pilot results confirm not only that marked performance differences occur within age groups as reported elsewhere (8, 49), but also between age groups. There is some suggestion that decline in performance sets in relatively early in middle age.

Keller's work with a mentally retarded boy suggests another experimental procedure involving not only recognition but also reproduction (16). Keller used Kohs block designs and requested his subject to reproduce the designs on paper which contained squares to represent the blocks. The stimulus was left before *S* while he made his copy. No time limit was imposed. *S* failed in his reproductive attempts, evidently being unable to come to the realization that the stimulus could be accurately reproduced on the squared paper by breaking the stimulus up into imaginary squares. In order to rule out the possibility that *S* was unable to discriminate between stimulus and reproduction he was presented with the test figure for five seconds and then required to select it from among four other figures, one of which was the correct match, and another of which was his own previous reproduction. Under these conditions *S* quickly and accurately selects the correct figure. *S* was next required to reproduce a test figure which was presented on squared paper identical with that used for the reproductions. *S*, despite this perceptual crutch, still fails. Initially, he was occasionally successful but as he proceeded with a design his reproductions suffered. Keller accounts for this behavior in terms of the "tyranny of the figure-as-a-whole." When *S* is required to reproduce diagonals with minimal configural properties he succeeds. As soon as the diagonals are joined in a configuration, *S* fails. In a final experiment *S* was provided with a frame which masked all but one square of the Kohs stimulus, and he was taught to go from square to square, building the design piecemeal. Keller reports:

> It was thought that this crutch might teach him that the total configuration could be periodically ignored and that the design could be built solely

> by careful attention to the individual units of which it was composed. He was able to operate very successfully with the frame, but when the frame was removed he reverted to his old confusions. Apparently, he could not by conscious control focus successfully on the parts so long as the whole was there to distract him (16, p. 564).

These findings are more striking when it is recognized that the entire procedure was conducted with therapeutic intent and every effort was made to help *S* improve his performance. These results reflect neither a simple perceptual defect nor a motor defect. The defect has been labelled a *specific cognitive defect* by Keller.

This kind of "functional analysis" is characteristic of much of the work of Kurt Goldstein and of the "brain-injury" researches of Heinz Werner. The specific techniques utilized by Keller are well suited for investigations designed to elicit performance differences in developmental and clinical samples and to determine the extent to which these differences are susceptible to modification by systematic training. The nature of such techniques permits inferential construction of central dynamisms which have predictive power.

The visual recognition techniques of Lowenfeld (21) with the modifications introduced by Wiggin open up still another path of potent exploration. Lowenfeld pulled line drawings behind a slot so that *S*s perceived only a limited aspect of the total configuration at any moment in time. *S* was then required to locate the figure among a group of figures presented for inspection. Wiggin standardized presentation by recording the stimuli on motion picture films (48). These techniques are applicable with both meaningful and nonsense materials.

For the entire series of visual recognition and reproduction experiments such factors as masking, interruption, time of exposure, context, etc., will be manipulated. The research program will supplement techniques previously discussed with those utilized by Werner, and Werner and Strauss with brain-injured children (39, 42, 45, 46, 47). These techniques address themselves nicely to the central research problem; namely, the specification of inferred central representational processes and the behavioral consequences of these processes for learning.

The basic design for research on learning, recognition, and reproduction in the visual modality is also applicable in the relatively unexplored field of stereognosis. A limited number of studies with animals (e.g., 34), in industrial psychology (15), in psychopathology (47), and in the area of creative activity have been conducted (21). These investigations contain promising techniques which can be modified to fit present needs.

Lowenfeld's work with tactual materials suggests a number of techniques useful in obtaining base performance rates on developmental samples. He presents *S*s with solid objects for tactual exploration and requires them to identify the stimulus object from among a group of objects. He also

used stencils cut in heavy cardboard. Woodworth reports a study by Yamane in which braille figures were utilized with blind and sighted-blindfolded *Ss* (50). Ruch, Fulton, and German used solid objects with primates under conditions of parietal intactness and parietal destruction (34). Werner and Strauss presented brain-injured *Ss* with relief figures and required their *Ss* to reproduce the stimulus (47). W. O. Jenkins employed various shapes and shape orientations in the course of developing a system of easily discriminable aircraft controls (15).

Revesz found rules for haptic perception akin to those obtaining in visual perception (33). There are very few developmental data on hand for this perceptual modality: collection of data will add information as well as serve theoretical ends. Modifications of available techniques will be utilized to obtain base performances and to test the assertions of association and cognition theories with respect to the specified variables as indicated by the basic design.

The haptic materials can be embedded, interrupted, varied in time, meaningful or nonsensical, varied in texture, pliability, and in size. Stereognosis may be an important area for exploration if its investigation adds essential data relevant to theorizing about learning and perception.

An adequate comparative psychology must deal not only with interspecies similarities and differences but also with intraspecies similarities and differences. General behavior theory has much to gain from the systematic collection of data within a developmental-comparative framework. These investigations may be fruitfully conducted with the aid of descriptive schemata and response inferred mediational mechanisms akin to those proposed in this paper. The value of such conceptions for biology is demonstrated by the notion of "organizing factors" introduced by the embryologist Paul Weiss (41), and more recently by the ontogenetic analogue of the course of nonspecific stress reactions in disease presented by Hans Selye (37).

In sum, this orientation permits specification of the properties of ongoing action and the relationship between ongoing action and acquisition variables. Predictions concerning the effects of inferred ongoing action properties are drawn from cognition theories. Since these predictions often are at variance with those of systematic learning theories, some clarification of this divergency is likely to emerge.

References

1. Aebli, H. *The development of intelligence in the child: summary of the works of Jean Piaget between 1936 and 1948.* Minneapolis: Institute of Child Welfare, Univer. of Minnesota, 1950.

2. Barker, R. G. Child psychology. *Annu. Rev. Psychol.*, 1951, 2, 1-22.

3. Beach, F. A. The snark was a boojum. *Amer. Psychologist,* 1950, 5, 115-124.

4. Dennis, W. Developmental theories. *Current trends in psychological theory.* Pittsburgh: Univer. of Pittsburgh Press, 1951. Pp. 1-20.

5. Friedman, H. Perceptual regression in schizophrenia: an hypothesis suggested by the use of the Rorschach. *J. genet. Psychol.,* 1952, 81, 63-98.

6. Gibson, Eleanor J. Improvement in perceptual judgments as a function of controlled practice or training. *Psychol. Bull.,* 1953, 50, 401-431.

7. Gollin, E. S. Developmental differences in visual recognition under conditions of stimulus impoverishment. In preparation.

8. Gollin, E. S., & Baron, A. Response consistency in perception and retention. *J. exp. Psychol.,* 1954, 47, 259-262.

9. Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.

10. Harlow, H. F. Learning theories. *Current trends in psychological theory.* Pittsburgh: Univer. of Pittsburgh Press, 1951.

11. Hebb, D. O. *The organization of behavior: a neuropsychological theory.* New York: Wiley, 1949.

12. Hebb, D. O. Drives and the c.n.s. (conceptual nervous system). *Psychol. Rev.,* 1955, 62, 243-254.

13. Hilgard, E. R. The role of learning in perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality.* New York: Ronald Press, 1951. Pp. 95-120.

14. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.

15. Jenkins, W. O. Tactual descrimination of shapes for coding aircraft-type controls. In P. M. Fitts (Ed.), *Psychological research on equipment design.* Washington, D.C.: U. S. Government Printing Office, 1947.

16. Keller, J. E. The investigation of a specific cognitive deficiency. *Amer. J. ment. Def.,* 1954, 58, 560-565.

17. Koch, H. L. Child psychology. *Annu. Rev. Psychol.,* 1954, 5, 1-26.

18. Koffka, K. *The growth of the mind.* New York: Harcourt, Brace, 1925.

19. Krech, D. Dynamic systems as open neurological systems. *Psychol. Rev.,* 1950, 57, 345-361.

20. LEWIN, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.

21. LOWENFELD, V. Tests for visual and haptical attitudes. *Am. J. Psychol.,* 1945, 58, 100-111.

22. LUCHINS, A. S. An evaluation of some current criticisms of gestalt psychological work on perception. *Psychol. Rev.,* 1951, 58, 69-95.

23. MCCANDLESS, B. R. Special review: recent books in child psychology. *Psychol. Bull.,* 1955, 52, 150-158.

24. MCGINNIS, E. The acquisition and interference of motor habits in young children. *Genet. Psychol. Monogr.,* 1929, 6, 203-311.

25. MILES, W. R. Age and human ability. *Psychol. Rev.,* 1933, 40, 387-414.

26. MILLER, G. A. *Language and communication.* New York: McGraw-Hill, 1951.

27. MURPHY, G. *Personality: a biosocial approach to origins and structure.* New York: Harper, 1947.

28. MURPHY, G., & HOCHBERG, J. Perceptual development: some tentative hypotheses. *Psychol. Rev.,* 1951, 38, 332-349.

29. PIAGET, J. *The psychology of intelligence.* New York: Harcourt, Brace, 1950.

30. PIAGET, J. *The origins of intelligence in children.* New York: International Universities Press, 1952.

31. POSTMAN, L. The experimental analysis of motivational factors in perception. In *Current theory and research in motivation: a symposium.* Lincoln: Univer. of Nebraska Press, 1953. Pp. 59-108.

32. RAZRAN, G. Conditioning and perception. *Psychol. Rev.,* 1955, 62, 83-95.

33. REVESZ, G. *Psychology and art of the blind.* New York: Longmans, Green, 1950.

34. RUCH, T. C., FULTON, J. F., & GERMAN, W. J. Sensory discrimination in monkey, chimpanzee, and man after lesions of the parietal lobe. *Arch. Neurol. Psychiat.,* 1938, 39, 919-937.

35. SCHEERER, M. Personality functioning and cognitive psychology. *J. Pers.,* 1953, 22, 1-16.

36. SEARS, R. R. A theoretical framework for personality and social behavior. *Amer. Psychologist,* 1951, 6, 476-483.

37. SELYE, H. Stress and disease. *Science,* 1955, 122, 625-631.

38. SIEGEL, E. L. Genetic parallels of perceptual structuralization in paranoid schizophrenia: analysis by means of the Rorschach technique. *J. proj. Tech.,* 1953, 17, 151-161.

39. Strauss, A. A., & Werner, H. Disorders of conceptual thinking in the brain-injured child. *J. nerv. ment. Dis.*, 1942, 96, 152-172.

40. Verplanck, W. S. Since learned behavior is innate, and vice versa, what now? *Psychol. Rev.*, 1955, 62, 139-144.

41. Weiss, P. *Principles of development: a textbook in experimental embryology.* New York: Holt, 1939.

42. Werner, H. Development of visuo-motor performance on the marble board tests in mentally retarded children. *J. genet. Psychol.*, 1944, 64, 269-279.

43. Werner, H. *Comparative psychology of mental development.* New York: Follett, 1948.

44. Werner, H. Developmental approach to general and clinical psychology. Presented as part of a symposium at Massachusetts Psychol. Ass., March, 1954.

45. Werner, H., & Strauss, A. A. Types of visuo-motor activity and their relation to low and high performance ages. *Proc. Amer. Ass. ment. Def.*, 1939, 44, 163-168.

46. Werner, H., & Strauss, A. A. Causal factors in low performance. *Amer. J. ment. Def.*, 1940, 45, 213-218.

47. Werner, H., & Strauss, A. A. Pathology of figure-background relations in the child. *J. abnorm. soc. Psychol.*, 1941, 36, 230-248.

48. Wiggin, R. G. A criticism of two of Lowenfeld's visual-haptic tests. Unpublished paper, Institute of Child Welfare, Univer. of Minnesota.

49. Witkin, H. Individual differences in ease of perception of embedded figures. *J. Pers.*, 1950, 19, 1-15.

50. Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* (Rev. Ed.) New York: Holt, 1954.

THE REPLACEMENT PROBLEM IN CHILD DEVELOPMENT

HAROLD E. JONES
University of California

Three decades have passed since the Laura Spelman Memorial began its program of subsidies to child development. The establishment of institutes, a generous fellowship program, and financial aid to journals produced a prompt increase in the number of persons annually entering the field, and correlatively a sharp rise in published research.

A part of this growth was normal, to be expected from the general growth of psychology and also from the earlier stimulus of the child study movement. A substantial part of it, however, was an outcome of forced feeding, and it may now be appropriate to inquire if this nutritional experiment was a sound and wise one.

The answer to this question may be a qualified affirmative—qualified in view of certain side effects which have not been wholly favorable to the normal growth of the field. One such effect was the almost inevitable counter-reaction of some persons at the core of the older disciplines, who felt that too much attention, and too much money, were being given to a new and untested branch of scientific effort.

In some psychology departments, steps were taken which to a degree weighed against graduate students with an interest in this field, and which were later to be used even more vigorously as a defense measure in dealing with a flood of clinical students. These sanctions sometimes took the form of graduate requirements heavily emphasizing historical and systematic psychology, and laboratory method. The prestige fields of psychology, closest in spirit to natural science, have always had some success in repressing offshoots. I do not wish to imply that there is anything malignant about this. Perhaps we should regard it chiefly as a kind of homeostatic control within the science, necessary to keep things in balance. Child development,

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

once in a stage of perhaps too rapid growth, encountered inhibitors, and some of these inhibitory mechanisms are still effective.

Another factor which needs to be considered is that the initial program of recruitment around 1930 brought into child development a number of persons who already had training and loyalties in other fields, marginal to science. It must be admitted that their interest in children did not necessarily imply a devotion to scientific method.

In the middle and late 1930's the artificial stimulation began to subside. A number of the centers originally dependent on foundations were undergoing a transition to local support. A fairly good balance was established between supply and demand of research workers and teachers in child development, and it was reasonable to expect growth—no longer at a pace of positive acceleration but of a slow and orderly nature.

World War II, however, brought new changes which were for the time being unfavorable to child development research. This field had little to offer in the prosecution of the war effort, or in dealing with veterans' problems after the war. In the national interest, very large resources were devoted to recruiting psychologists into clinical work, and also into various aspects of military psychology. Of all the branches of psychology, Developmental has undoubtedly suffered the most from the subsidizing of clinical training. Many graduate students who would otherwise have worked in child research, and whose interests were not primarily in therapy, could hardly be expected to resist the stipends and the superior resources for training which were made possible, on a grand scale, by governmental agencies.

Since 1952 somewhat greater funds have become available for fellowships and grants-in-aid in child development. The National Institute of Mental Health has had a liberal policy in this regard, and we now know that many aspects of child development, especially on the biological side, are relevant to the program of the National Science Foundation.

Although no foundation has taken the place of the Laura Spelman Memorial, in a primary concern with children, we can no longer feel that the sun is under a cloud so far as research support is concerned. There remain however many problems in our recruiting program. From the standpoint of the recruit, what kind of a field will he or she be entering? Is it one which presents a cheerful front to the world, with some pride in the past and confidence in the future, or is it one marked by self-doubts and querulous complaints?

After a recent visit to this country, Dr. Skard of the University of Oslo reported on Recent Trends in Child Psychology in U.S.A., in *Courrier*. Her opening statement is that, "The most striking impression that research in child psychology gives today is one of great differences and divergencies in problems and methods. A latent dissatisfaction seems to be present among the scientists in this field." Now, this is much better than accusing us of complacency. But dissatisfaction may be of several kinds. It may be con-

structive or it may be destructive. A lively interchange of criticism of specific studies may help to improve methods and lead to new answers to new problems. But if there is a general melancholia about the field, or a cynical attitude about all approaches except one's own, we can hardly expect to present an attractive picture to students.

Instead of deploring divergencies, I think we should regard individual differences in the *modus operandi* of research as a very favorable sign. We should recognize that psychology, and child development, flourish on individual differences. It is, I think, extremely fortunate that we have vigorous interests in laboratory research on theoretically loaded problems. It is also of great importance for the future of the science that we have an expanding interest in theoretically neutral research, and in the patient application of exploratory observational methods. There is a bright promise in cross-cultural research, and in the renewed application of comparative or cross-species methods in studying some genetic problems. We should welcome the fact that we can shelter and provide expression for a great variety of emphases in clinical, quantitative, descriptive, interpretive, naturalistic, experimental, and longitudinal approaches.

But it is a somewhat puzzling fact that a few persons who are happy in only a limited corner of the field seem unhappy that other areas exist. Their efforts to devalue what they are not themselves doing may result merely in devaluing the whole field. This situation has been examined earlier in this conference, with regard to the strategy of research, but my concern with it here is in connection with the effect upon recruitment.

Of course, when we consider the actual number of persons, among psychologists, who are working with children we find that our clientele is by no means a vanishing one. In 1918 there were only 3 psychologists who indicated a primary interest in children. By 1937 this number had increased to about 81. At the present time nearly a thousand members and fellows of the APA indicate the child, or work with children, as their first interest. It is however true that too few of these are active in research, and that measures are needed to increase the training of research personnel.

Proposed measures may be considered first at the national level and then in terms of what might be done in individual institutions. Obviously the most important factor in recruiting students is to have something to recruit them into: an active and healthy research field, and one which is not too much under the influence of the need for aggression or the need for abasement. In order to compete with other fields which are already well supplied with stipends, I would propose a national program of competitive fellowships specifically for child development research. We might appropriately discuss how such a program could be set up, and under what auspices. Even if it included no more than five or ten awards a year, it would have a value going considerably beyond the actual number of appointments. I would also advocate a national committee to deal with problems of child

development teaching, to administer visiting appointments, and to study instructional programs, instructional aids, and graduate training curricula. This cannot very well be done by an *ad hoc* committee appointed by a society. It should have funds, and representation from different societies.

Finally, the effect of such measures should be to extend the number of institutions which, like Minnesota, Chicago and Iowa, have a training program specifically in human growth and development. Where a student interested in child development must take his degree in a psychology department, he sometimes finds that the demands of the "classical revival" make it all but impossible for him to conduct a well-rounded program. He is well trained in psychological systems and conventional laboratory methods, but finds it difficult to get what he needs in physical growth, developmental physiology, or in courses in sociology or anthropology bearing on child development. Lacking time to conduct a field study, he may find that the line of least resistance toward a Ph.D. is through the rat laboratory or in the customary exploitation of college sophomores.

Every now and then I like to refer back to a statement made more than a quarter century ago by W. I. Thomas, in reference to the genetic method, which as he said represented "a reaction against the practice of observing behavior at a given moment by a given technique (e.g. psychological, observational) and drawing inferences as to causation. It is realized that the conception can no longer be entertained that it is adequate to correlate any two variables—a given behavior manifestation with nutrition, or skeletal development, or mentality, or birth trauma, or family, or bad company, taken separately. The genetic method proposes, therefore, to study human growth as a maturational and experiential process from early infancy through adolescence and beyond, in individuals, groups and control groups, with reference to both physical maturation and the complex of cultural factors (family, school, moral code, etc.), by the simultaneous and continuous employment of all available techniques—anatomical, physiological, psychological, etc."

For a generation ago, and perhaps for the present generation also, this may be too ambitious a statement. Our child development research must often be contented to whittle away at some parts of this interdisciplinary programming. We must make many compromises. But we should make them reluctantly rather than triumphantly, and we can hardly forget that if we are really interested in child development we must sooner or later come to grips with changes through time, and not merely with the horizontal cross section. Graduate programs which are aimed only at cross section or only at situational research may not be well adapted to recruiting and training all of the kinds of persons needed in this field.

THE NEED FOR CONCEPTUALIZATION IN RESEARCH ON CHILD DEVELOPMENT

IRVING E. SIGEL
Merrill-Palmer School

To be asked to pose research priorities in child development presupposes that we are dealing with a contained discipline in which basic issues and problems are clear. Child development, however, at this time does not have a coherence, and consequently a statement of research priorities emanates from idiosyncratic perceptions of the field rather than from deductions derived from the conceptual unity of the field.

In order, then, to discuss priorities, it seems that we must come to grips with a prior question, namely, what is child development? The answer to this question is a necessary step in posing priorities in research, for ideally they can flow from the problems residing within the area.

The question, what is child development, is old. Today, it has various connotations. To some it is merely an interest area; to others it is a point of view; to others it is an area with a specific content; to still others it is an empirical hodgepodge of unrelated facts and speculations dealing with the child.

For me, child development is an aspect of human development. The emphasis on the childhood period is partly idiosyncratic and partly based on the assumption that the basic patterns of many human characteristics and processes have their beginnings here. The goal then of child development is to understand the processes and content of the "progressive" change of the organism, specifically differentiation, specialization and integration of human structures and functions as the organism moves along the continuum of the life cycle. The functional unity of child development and of its broader aspect, human development, stems from the common integrating principle of development. The developmental set is crucial in distinguishing child development from other positions studying the child.

If the premise is accepted, then some of the basic research questions facing us, derived from the developmental orientation, are as follows: what changes occur within the organism as he matures, what conditions affect the course of this change, what processes are involved in change, and how these changes influence and are influenced in the organism's interacting and coping with his environment.

With the task so defined, the question becomes, how do we go about it? The purpose of this paper is to propose a procedure for so doing in order to facilitate the achievement of the goals stated above.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

If a segment of human knowledge is to be organized under a single rubric for the purposes of scientific study, it behooves us to work towards the establishment of a conceptual framework within which a theory or theories can be evolved. Before proceeding with such a comprehensive task, we should ask ourselves if there is any current conceptual orientation which attempts to integrate knowledge from various established disciplines into a comprehensive unit.

The "holistic approach" is one point of view which is prevalent among child development workers. It holds that we are concerned with the total organism, that the organism in its unity is different from the sum of its parts. If this assumption is correct, the course of action to be followed should stem with methodological consistency from such a conceptual position. It requires that methodologies be set up so that the whole can be studied simultaneously in all its complexity. To my knowledge this has never been done, nor at this point in our development, do I believe it can be done.

Whether the organism is such a totality functioning in such unity is still an open question. The acceptance then of the assumption precludes the investigation of relationships of aspects of the organism, since this, too, does violence to the essential nature of the individual.

Further, this position leads to the conclusion that our existing knowledge about human development is in error, since little if any of it was gathered under holistic conditions. The organism has been studied at different time intervals and in its different functions. Yet we know that we have bodies of verifiable knowledge useful in predictions of its future reactions.

Thus I see this holistic orientation as being inconsistent, at least in regard to theory and practice. For a unified science we need theory consistent with practice. This then means that the empirical research within a discipline should ideally be compatible with the theoretical position or consistent with the conceptual framework taken.

The holistic position, to me, is a philosophical one in which highly abstract concepts can be used to integrate the totality of man in his universe. But to presume that a body of empirical knowledge can emanate under such a point of view seems highly doubtful. If we leave the holistic approach to the realm of philosophy, I believe we will be in a better position in evolving a conceptual framework which can relate coherently to empirical research.

Rather, I would prefer to think of the "whole" child as an organization of a number of systems, which function at varying degrees of autonomy and interrelatedness.

The first step in organizing such a conceptual framework is the identification and delineation of systems within which we can organize coherent developmental material. Next we need to gather together existing data describing the course of development within the designated systems. After this, we should aim to describe the stages of development within systems and between systems in order to describe the concomitant appearance of struc-

tures and functions; with this achievement we can proceed to examine relationships between systems to determine the degree, the kind, and the quality of relationships existing between systems.[1,2]

In setting about to delineate and define systems of the organism we become impressed with the unalterable fact that the course of a child's development is a function of a variety of variables residing primarily within the province of biological and social sciences. Thus the science of development seeks relevant variables in such disciplines as biology, psychology, anthropology, to list but some. The need to be dependent upon many disciplines in our construction of a single conceptual framework stems naturally from the nature of our problem.

We are immediately faced with the question, *what* do we incorporate from the various disciplines? Do we take concepts, or facts? Do we take theories? Do we take whatever we deem necessary? If so, what are the criteria for this necessity?

Depending on whether we incorporate theory, concepts, assumptions or facts, we face different problems. When we borrow concepts we must ascertain their meaning and their utility for answering our own needs. Is the concept borrowed relevant for our material? For example, homeostatsis is a case in point. It is used in physiology and in psychology. Is its meaning equivalent in both instances? Or are we just borrowing words, which we then redefine to meet specific conceptual questions? Or are we attempting to fit a concept onto a set of data which has been derived from very different sources? I feel these questions must be posed when we attempt integration of any kind. The flight into someones else's conceptual structure may be more impeding than facilitating.

We can ask similar questions when we attempt to incorporate theoretical positions. Is it a flight into theory because that theory has direct meaningful relationships and utility for our purposes? Or is it rather an expediency? In other words, do we try to set the theory onto a given set of empirical facts or does the theory emanate and develop from various observations and interpretations of observed phenomena? Again relevance, utility and appropriateness should be used as guides. One more point here is that we should also examine the assumptions upon which a theory is based. Such assumptions, altogether too frequently are not explicitly stated, and may be inherently incompatible with that of the borrower.

When we seek to incorporate facts from other disciplines, we also need to be cautious. The usual criteria employed in assessing any scientific fact

---

[1] To be sure, some steps have been taken already in the directions cited above. Baldwin's (1) recent text is a notable example of the application of psychological concepts to development. More concentrated efforts should be expended in these directions in the next few years.

[2] Although these stages are set in sequence, there is great overlapping. Thus, investigation of development within a designated system may also yield data on relations between systems. Thus the order here appears more rigid than intended.

should be employed. A fact obtained, however, under one conceptual system may not be a fact under another system. For example, if the corroborated findings of experimental research indicate that children from authoritarian settings are more punitive than children from nonauthoritarian backgrounds, we must be clear as to what punitiveness is. Facts then will have to be evaluated from the context of their origin.

Assimilating concepts, facts or theories from other disciplines does not mean, however, that child development has to be a receptive science—receiving contributions from others only. In the initial stages of organization and integration such receptive dependence may be necessary. As child development progresses in its unity, it will evolve more and more its own body of theory and fact, and hence move toward greater autonomy. However, it should be recognized that some interdependence of all the sciences exists. Let is not be thought that I conceive of child development as progressing toward insulation or isolation. Rather I conceive of it as developing in its unique way, facing its unique problems and so contributing to the study of man. In so doing it can furnish other disciplines concepts, theories and facts.

With some of these considerations in mind, the task facing us is to delineate the systems and define their content. Admittedly such a task is somewhat arbitrary, but for the purposes of study and research this approach is necessary. Let it be remembered, however, that the arbitrariness with which this is done is utilitarian and relative. This is no plea for absolutism.

We start with the premise that the developing organism in its environment comprises our field of interest. How we choose to delineate the systems depends on the level of abstraction at which we elect to start and the nature of the concepts we elect as guides. For example, we can start with such distinctions as biological, psychological and social. Or we can start with the personality and the environment. Or we can proceed by viewing the total in a functional-structural dichotomy and proceed from there. The particular starting point derives from our basic conception of human nature and of our perception of the organization of knowledge.

A starting point for now may be a tripartite breakdown of biological, psychological and exogenous environmental systems. Within each of these further delineation is possible, e.g., within the biological we can separate skeletal, or neural subsystems; within the psychological such subsystems as cognition, perception can be distinguished. Further identification of smaller units within some of these second order categories can be isolated.

Hopefully, the concern with each of these systems would be with process and with content. Thus, the interest, in psychological systems, for example, would be with the processes, e.g., cognitive process, and with the cognitive content, e.g., attitudes, beliefs.

I should like to digress here momentarily to emphasize one of the three large systems mentioned above—the exogenous environment—since the con-

sideration of the exogenous environment as a functional variable has received little systematic attention. In view of this rather serious omission, I should prefer to discuss this at some length, fully aware that similar elaborations are necessary for other factors.

As Chein (3) has pointed out, the environment must be systematically schematized so that appropriate representation of variables within the environment can be differentiated. The environment can be analyzed structurally, e.g., social structure, social class, family organization. What relationship exists between the child and various aspects of his environment? If we conceive of the environment as geo-behavioral, that is, the properties of the physical-personal environment as responded to by the reacting individual, then we can conceptualize the environment in functional terms, e.g., supporting, constraining, or directing the development of the individual.[3]

From this point of view, then, we can investigate the differential influence of various environmental experiences upon the developing child. As the child matures, certain aspects of the environment may function differentially; e.g., the home environment, considered globally for illustrative purposes, may have a high frequency of supportive and directive points of contact during infancy, whereas in later childhood this may shift to greater constraint and less support.

Such emphasis on the environment is especially important in the study of the socialization process. Extensive description of various points of environmental contact would give us greater insight into the manifold experiences which are involved in the child's socialization. Barker and Wright (2) have taken some significant steps in this direction and it is necessary to extend such investigations into various other types of communities. This is especially true for children living in urban areas.

Once we have determined the relevant developmental systems we need to describe adequately the developmental course for each of them. What are the developmental trends in various aspects of cognitive development? Piaget (5), for example, has given detailed description of some of the stages of development in this area. Gesell (4), Shirley (6) and others have presented descriptive material on physical growth trends. Additional descriptive data are needed. Where the need is can be most readily ascertained once we begin some sort of organization of the separate systems and assess the state of knowledge within them. This procedure is based on the premise that adequate description of events is a necessary phase of growth in any science.

To be sure, we are not content with descriptive data, for we do not conceive of our science primarily in taxonomic terms. What we seek are

[3] The terms are those used by Chein, but the definition of the environment is not identical to his.

explanatory principles and theoretical constructs which will further our understanding of developmental phenomena.

In the search for explanations and for laws of development, we are forced to examine the relationships between systems as well as the presumed homogeneity within systems. To what degree then is a specific system homogeneous? And further, to what degree is a specific system related to other systems?

As we proceed to answer these questions, we may arrive at clusters of systems which are presumably more interdependent than others. These clusters I like to call self-contained units. Let me illustrate. Take three systems: social structure, adult personality, and child rearing practice. We can examine each of these separately, if we seek answers to such questions as, what kind of social structure exists in a certain society, how does this society function, etc. Or if we are interested in personality or child rearing in and of themselves we can investigate each of these as a separate.

If, on the other hand, we choose to study how social structure influences adult personality, we may arrive at the notion that it does so primarily through its child rearing practice. Hence we can examine the relationship among these three—a self-contained unit—without involving any other system. The nature of the questions posed, then, becomes one of the determinants of the organization of self-contained units.

Yet it may well be that there exist "natural" self-contained units, within which systems are more intimately tied. Our empirical research, testing our conceptual organization, will indicate the relative closeness of systems.

A further step is to conceptualize and test the relationship between self-contained units. The implication here is that even self-contained units may not be completely autonomous but may have only relative autonomy. By research derived from the type of conceptualization proposed here, we can determine the degree to which self-contained units are closed and have no relevant dependence, under certain conditions and to limited degrees, on other units.

It should be said here that the concept of self-contained units is not synonymous with the concept of a system. A system is an assemblage of processes or structures collectively contributing toward specific functions or processes. Thus, processes and structure of cognition, for example, can be viewed as a system. A self-contained unit may contain one or more systems. The criterion of belongingness is the degree of mutual interdependence. Thus, physical growth can be viewed as a self-contained unit comprising numerous systems, e.g., skeletal, neural, circulatory, etc. The concept self-contained units is heuristic at this time. In time, as our conceptualization improves, such heuristic concepts may prove to have theoretical value.

As soon as we begin to examine the interrelatedness of systems and/or self-contained units, particularly if these are derived from different disciplines, we are faced with problems of conceptual and factual comparability

and equivalence. In the social structure—modal adult personality—child rearing question, for example, we seem to be dealing with extensive and abstractive concepts which appear equivalent. Thus, we would separate out aspects of each of these and indicate the relationships so that we could see what elements of social structure affect what elements of personality. Further, we would have to find out what aspects of personality affect what child rearing practices; and finally, what differential effects various child rearing practices have on the developing personality of the child. In this way, presumably lawful relationships can be empirically demonstrated so that our abstract generalizations about the social structure—adult modal personality—child rearing practices chain reaction can be made with some degree of specificity.

If, on the other hand, we do not tend to see conceptual equivalents and the variables differ in their level of complexity, without considering these, then we do not know what affects what. For example, when we seek relationships between physical and mental characteristics, we find low relationships. This may be due to the fact that the variables contain different degrees of complexity so that the correlations are low, not because *no* relationship exists, but rather elements within the variables not identified or made explicit are included which operate to reduce the relatedness. IQ and skeletal growth studies are good illustrations of this. The various complexities represented in the single IQ score may mask or counterbalance each other, or be unrelated. Thus, low or insignificant relationships are found.

A further problem that occurs here from the developmental framework is that at various stages different relationships between variables may exist. For example, in the first year of life, motor skill and mental ability are more closely interdependent, and seemingly causally interrelated, than at age ten. The infant needs certain motor skills to handle certain intellectual tasks, whereas at an older age, with language, motor skills assume a different role.

Thus, when we view the relationships between systems of the organism developmentally, we may need to indicate the relationships sequentially as well as concomitantly. Some systems, such as motor skill or physical growth, have a stage setting function—they prepare the organism for performing certain tasks, while other systems may be correlated in their appearance in time, such as fine motor skills and the beginnings of conceptual thought.

The schema presented for systematization of the material within the child development area is admittedly broad and extensive. The task of ordering the material into meaningful units requires a more detailed and specific conceptual frame-work than was described here. This I feel is a research task and is one which hopefully can proceed in the next few years.

If such a goal is deemed important, then it might well act as a guide to our selection of problems, developing of concepts and interpretation of

empirical fact. As we proceed, then, in such a direction, we continue to strengthen the scientific basis for the field.

One practical approach to such systematization can be in our graduate teaching program. Graduate seminars can be organized around conceptual problems where efforts are directed toward definition and delineation of systems, integrating existing information within these systems and possibly developing concepts for hypothesizing relationships within and between systems. Hopefully, a task of this nature could instigate related research for dissertations.

Sharing of this type of teaching experience by articles in professional journals or discussions at appropriate national meetings would be of real value for interested workers in the field. The sharing of such experiences should be of value in furthering our joint efforts in such a direction.

The one research area I would like to focus on in the remaining section of the paper deals with the nuclear family's impact on the personality development of the child.

As the primal social environment of the developing child, the intensive study of nuclear family interaction and interrelationships holds the key to much of our understanding of the developing child, particularly in the earliest years of growth. Although much research has been done, we need to know much more than we do about the complex involvements, interactions and relationships of family members upon each other through the life cycle. In so doing we can ascertain the factors influential in affecting the course of personality development.

The family is viewed as one complex environmental variable having the highest frequency of points of contact, at least in the early years of the child's life. Some major questions then arise in our attempts to describe these contacts, to indicate their effect on the child at varying points in his development. In Chein's (3) terms, for example, analysis of the family, in its structure-function organization as an environmental variable supporting, constraining or directing the child's growth, should yield valuable data as to how and in what ways the family influences the child's developing personality.

The steps in this programmatic research suggestion are similar to what was mentioned previously about the whole field.

1. What is needed is increased descriptive data on the functions of various family members as they interact at various stages of the individual's development. In previous research in parent-child relationships, we have dwelt too heavily on the mother-child relationship. Important as this relationship is for the child, let us not ignore the male member of the family. Ironically, in our Freudian psychology the father plays a prominent role, yet in our actual research in parent-child relationships we have tended to ignore the role of the father or to study his role reconstructively, using mothers and children as informants. Thus, research should be concerned with the

father's role in his relationship with his child in its many aspects. Such problems then emerge, as the type and quality of contacts the father has with his children, the differential effects of such experience for the father and how these change with changes in the child. The mother's role should of course be examined in a comparable fashion. Much, of course, has been done as far as the mother is concerned. Both parents should be necessarily involved.

The next step here is to study comparatively and interactively fathers and mothers in terms of relative effect and influence. Thus, we are interested in the relative influence of parents in control of the family and of the children. What is the effect on the developing child of inconsistency between parents? Further, what difference does it make who is the frustrator, the gratifier, the indifferent one in relation to the child's need satisfactions? In other words, we have to focus on the parents comparatively to assess and evaluate the type and kind of points of contact they have with the child.

2. There is also great need to study the other members of the nuclear family, especially siblings. The interactions that occur within the family that have potential significance for the child stem from siblings as well as from parents. The siblings in their relationship to each other may reinforce, positively or negatively, parental actions, and consequently may function to maintain or to change potential responses of a child to his parents. For example, when an older sibling functions to gratify the love needs of a younger sibling, while the parent is frustrating such gratification, would we not expect different personality characteristics to appear as compared to the situation where no sibling performs the love-giving function?

The analysis of the family as it functions in various ways to effect the child should follow a longitudinal pattern in order to assess its role at different developmental stages of the child. As the child matures and his familial relationships begin to function differently, the influence of the nuclear family may vary. Thus, we have to view the two variables here, the child and the family, as two changing systems whose interactive and influence patterns will shift in relation to different demands and needs of each. As we follow this trend, seeking specific functional relationships, we should be able to specify in necessary detail those variables most readily influenced by familial experiences.

Families in such research should be drawn so as to be more representative of the population as a whole. We need to devise ways to contact and to maintain rapport for research purposes with lower income families, with families of different ethnic groups, and in general, nonclinical populations of families, in order to get the variations of family functions and their different influences on the child.

The more intensive our study of the family, and the sets of conditions residing therein, the greater should be our knowledge of the individual differences of developing children.

The relationships that exist within the family are to a large measure expressed in the type of interactive behaviors that occur among family members. Feelings of love, respect, trust, or the negative aspects of these take some concrete behavioral form. Then it behooves us, as one point of emphasis, to study the effect of these overt behaviors. What family members say and do to each other in the course of their living together should be evaluated in terms of impact on the family participants.

Within this context of behavioral investigation I should like to emphasize the urgent need for intensive study of the role of parent behaviors in influencing the numerous aspects of the child's development. Whereas we at Merrill-Palmer are focusing on parental control behavior, there are many other categories of parent behavior that are relevant, e.g., play, social interaction, etc.

Thus, we select out one aspect of the family environment, the behavioral one, and must set out to describe parent behavior in relation to various experiential conditions, e.g., what does the parent do when the child comes for comforting, or what does the parent do when the child is frustrated? But we must also be concerned with how the parent changes in behavioral interaction with the child as the child gets older, when the issues and needs about which parent-child interactions occur may take more complex forms.

In our study of parental behavior, we have many complex theoretical and methodological problems. A major one seems to me to revolve about the effective stimulus variables within the behavioral act. What is it that stimulates the child to respond? For example, what is a joking statement and said in a joking tone with joking words may be in reality a threatening anxiety-producing stimulus to the child. Is it the fact of joking, is it the timing of the statement, is it the way the parent looks, or what?

In examining such questions as to the effect of the stimulus, it seems that we must include as intervening variables the child's perceptual and cognitive functions. Thus, the study of the behavior of parents toward their children points to the development of a theory in which such systems are interrelated. This includes the acts of parents, the responsive process in the child, and the structure of the stimulus itself. We will have to integrate concepts from various psychological subsystems, e.g., personality, cognition, etc., in order to explain adequately the role such behaviors play at different developmental stages.

In addition to focusing on the behavior of the parents, we must not ignore the influence of the child's behavior on the parent's behavior. In other words, this interaction is a mutual one and hence some of the contemporaneous antecedents of each set of behaviors lie within the behaviors of the other person. Thus, the child's acts stimulate the parent and the parent's acts stimulate the child. When we look toward the role child behaviors have in influencing the parents, we again need to ascertain the meaning these child behaviors have for the parent and what properties of

child behavior instigate various adult responses. Here we become involved in the value system, the motivational system and the cognitive system of the parent, as antecedents for what child behavior will be response producing and for the nature of the parental response.

Up to this point we have examined the family from within and how its members influence and affect each other. But we can and should study the family as a unit and its relationship to its larger environment. Since our society is an ever-changing thing, since it is true that the growing child in his family is dynamically changing, then we are faced with a research problem involving the relationships of two dynamically changing systems. Here we are interested in the effect of these two variables on each other. What is the effect of a changing society with shifting values, with complex requirements for achievement, and with undefined customs and mores, on the developing child and his family? How do these societal forces influence the child's adjustment, what and how many adjustments does a child have to make, what happens to the child and the family in their sets of expectations of behavior from others, to achievement in life, to peer and other social relationships? Finally, what is the effect of these societal instabilities on the relationship between children and their parents?

Within this complex society, we have to examine the influence on the child and his family of other individuals, other relationships. One microcosm that is of importance in studying the factors involved is the neighborhood. This is particularly true in urban areas, where this geographical unit is to a large measure the habitat of the young child. Once the child starts to move in it, people other than family members are going to play a part in influencing him directly or indirectly. His experiences in these situations may bring new problems into the home which were never anticipated by parents. Although these may seem at this point spurious and unpredictable forces impinging on the child and his family, the investigation of the world of the child to include these variables would aid in throwing some light on the role of ecological forces in the development of the child.

I would prefer this type of approach at this time. We frequently talk of cultural forces when we speak of the influence of the social environment on the child. This is a very gross point of view, particulaly in a complex and heterogeneous society as ours. We have to ask the question, what unit of culture are we to study? In our society, culture and social class are extensive abstractions and necessary concepts in the broad social picture of examining human growth and development. Neighborhood is a utilitarian concept. No doubt it reflects many of the broad cultural patterns, but at the same time it selects and shapes these into its own image and it is this unique microcosm to which the child is making some of his early adaptations. For example, materialistic values may be said to be dominant in this country, but it is the interpretation of these in a particular neighborhood and the social consequences of the locally approved pattern that is having its impact

on the children under study. Thus, by focusing on field studies which specifically deal with home-neighborhood-school constellations we gather the specifics which are necessary in order to make our adequate generalizations. An excellent example of the kind of approach I am advocating here is the work of Barker and Wright (2) in their studies of the ecological habitat.

These questions and problems, although briefly stated are obviously extensive in scope. Yet I feel a many-pronged attack on various aspects of them is of vital importance in our understanding of the development of the child within his psychosocial environment.

The practical and theoretical contributions made by such study are numerous. For as we improve our understanding of the intricacies of the developmental process within the family and its social environment, it should help us to understand the patterns that exist, the deviations that occur and eventually the therapeutics needed to maintain a healthy course of change and growth of the child and his family.

References

1. Baldwin, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.
2. Barker, R. G., & Wright, H. F. *Midwest and its children.* Evanston, Ill.: Row, Peterson, 1955.
3. Chein, I. The environment as a determinant of behavior. *J. soc. Psychol.,* 1954, 39, 115-127.
4. Gesell, A., et al. *The first five years of life: A guide to the study of the pre-school child.* New York: Harper, 1940.
5. Piaget, J. *Language and thought.* New York: Humanities Press, 1951.
6. Shirley, M. M. *The first two years: a study of twenty-five babies. Vol. I. Postural and locomotor development.* Minneapolis: Univer. of Minnesota Press, 1931.

EXPERIMENTS WITH CHILDREN ON THE HYPOTHESES OF ACQUIRED DISTINCTIVENESS AND EQUIVALENCE OF CUES

CHARLES C. SPIKER
Iowa Child Welfare Research Station

The participants in this conference have been asked to concern themselves with matters of high priority with respect to research in the field of child development. This writer takes the position that research problems lending themselves to integration within existing theory, or which contribute directly toward the construction of theory, are among the most fruitful problems that can be investigated at any given time. Since the writer views the field of child development as a multi-disciplinary field, since he is convinced that few scientists are genuinely competent in more than one or possibly two disciplines, and since his own training has been in the psychological area, he prefers to restrict the range of his comments to the area of child psychology.

It does not seem that the child can profitably be viewed as a unique subject for psychological investigation. Rather, the child is another subject for psychological study, and, for theoretical and/or practical reasons, certain problems are more appropriately studied with children than with other subjects. The results of such research with children are thus expected to contribute to a general theory of behavior and not to theories of child behavior. This conception does not deny that productive research workers with children must have some formalized training in the handling of children. But it would argue that however important, socially and morally, we may consider such training, it does not differ fundamentally from the type of training that is provided the student of rat behavior. In particular, the necessity for such training does not vitiate the equally important need

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

for training in the areas of the philosophy of science, research methodology, experimental design, statistical analysis, and general psychology.

The remainder of this paper is largely concerned with an illustration of the manner in which psychological research with children can contribute to general behavioral theory. It will present a review of some recent research with children dealing with two specific hypotheses from behavior theory.

The first of these has been variously referred to as the hypothesis of secondary generalization (14), mediated generalization (3), or the acquired equivalence of cues (18). In brief, it states that if a subject has been trained to make the same (mediating) response to two or more dissimilar stimuli, there will be an increased tendency to generalize to the other stimuli other (mediated) responses *subsequently* learned to one of the stimuli. This hypothesis is derived from familiar conditioning principles plus some specific assumptions concerning the occurrence and function of response-produced stimulation. The derivation may be sketched as follows. Learning the mediating response (R_1) to each of several dissimilar stimuli (S_1, S_2, . . . S_n) assures that on subsequent presentation of these stimuli, the stimulation (s_1) produced by R_1 will be present. Any other response (R_o) subsequently learned to one of these stimuli (S_1) will also be conditioned to s_1. Thus, if any other of the stimuli (S_2 . . . S_n) is presented, R_1 will occur, together with s_1, and s_1 will tend to evoke R_o.

The second hypothesis has been referred to as the principle of acquired distinctiveness of cues (18). As the name implies, this hypothesis has to do with discrimination rather than with generalization, although the basic assumptions involved in the two hypotheses are highly similar. Briefly, the hypothesis is that if a subject learns to respond differentially to two or more similar stimuli, he will subsequently learn more rapidly other discriminal responses to these stimuli. The learning of the first set of discriminal responses (R_1, R_2, . . . R_n) to the stimuli (S_1, S_2, . . . S_n) assures that on subsequent presentations of these stimuli, the stimulation (s_1, s_2, . . . s_n) produced by the discriminal responses will be present. Hence, when S_1 is presented, the total stimulus complex will include s_1; when S_2 is presented, the total complex will include s_2; and so on. Assuming that s_1 is more discriminable from s_2 than S_1 is from S_2, the learning of the first set of discriminal responses should serve to reduce the functional similarity among the stimulus complexes. One of the implications of this set of assumptions is that a new set of discriminal responses (R_a, R_b, . . . R_z) could be learned more readily to the stimulus set (S_1, S_2, . . . S_n) if the subjects have first learned other discriminal responses than if they have not. This implication follows from the familiar conditioning law that relates the difficulty of a discrimination learning task to the similarity of the discriminal stimuli.

Some of the more convincing empirical evidence for these hypotheses has been obtained with children. One of the earliest investigations of the

hypothesis of acquired equivalence of cues (AEC) was conducted by Birge (3). Her grammar school subjects first learned to call two of four nonsense animals by one name and the other two by another name. One stimulus from each name-pair were presented together and the subjects learned to find a piece of candy under one of them. Finally, the other member of each name-pair were presented together and the subjects were allowed to choose one of them. She found that the majority of subjects chose that stimulus which had the same name as the stimulus under which they had previously found candy. Moreover, the tendency to choose in this direction was greater for subjects who were required to verbalize the names overtly during the candy-finding and test stages than for subjects not required to verbalize overtly during these stages. In a further experiment, Birge found out that subjects were more likely to generalize a name from one stimulus to another if they had found candy associated with both stimuli than if candy had been associated with only one of them. Shepard (24) replicated a part of Birge's first experiment, using different stimuli, and found essentially the same results.

An experiment by Jeffrey (15) with preschool children provides another demonstration of the role of mediating responses in generalization. He trained his subjects to move a lever in one direction to a white stimulus and in the opposite direction to a black stimulus. He then taught some subjects to call a gray stimulus "white" and others to call it "black," and next retrained the subjects on the lever moving task to the white and black stimuli. Following this, he presented the gray stimulus interspersed with black and white stimului to determine how the subjects would respond to the gray stimulus. He found that if the subjects had been taught to call the gray stimulus "white," they responded to it as they did to the white stimulus; if they had been taught to call the gray stimulus "black," they responded to it as they did to the black stimulus. He found a similar tendency, though not so strong, for another group of subjects who instead of using the names "black" and "white," had been taught to turn the handle to the right or to the left for the white, black, and gray stimuli. Thus, like Birge's experiments, Jeffrey's indicates that either verbal or motor (manipulative) responses may serve as mediating responses.

The generality of the AEC hypothesis is suggested in a recent experiment by Eisman (9) in which she produced color preferences in children through mediated generalization. She first taught grammar school children a different name for each of three white geometric blocks. She then reinforced them for selecting a particular one of the three blocks. Next, she taught the children the same three names for three differently colored blocks. When the subjects were then asked to find the reward under one of the colored blocks, they tended to choose that color which had been assigned the same name as the rewarded geometric block. By giving addi-

tional naming experience with colored jar tops and glasses, she was able to demonstrate a selective preference for that one of three white blocks that had been assigned the preferred color name, and also a verbalized preference for that one of three fictitious groups of children that had been assigned the preferred color name.

Results consistent with the AEC hypothesis have also been reported with adult subjects by Murdock (19), Bugelski and Scharlock (4), and Russell and Storms (23). In the latter two studies, it was reported that the subjects were unable to verbalize the basis of mediation.

The empirical evidence for the hypothesis of acquired distinctiveness of cues (ADC) is somewhat more difficult to interpret. The simplest paradigm involves two groups of subjects given a pretraining and a transfer task. The two groups are given the same transfer task under identical conditions, but differ with respect to the type of pretraining. The experimental (relevant stimulus) group is given preliminary training in which distinctive responses (usually names) are learned for the stimuli involved in the transfer task. The control (irrelevant stimulus) group is given similar training with stimuli that are not involved in the transfer task. The transfer task itself requires both groups to make other discriminal responses to the stimuli. This paradigm is assumed to control for positive transfer attributed to such non-specific factors as "warm-up" or "learning to learn." Thus, the relevant stimulus group is expected to learn the transfer task more rapidly than the irrelevant stimulus group due to the contribution of the response-produced stimulation to the distinctiveness of the total stimulus complexes.

The results of research with adults on the hypothesis have been conflicting. Arnoult (2), J. Cantor (6), Gerjuoy and Bechtoldt (12), McAllister (17), and Robinson (21) were unable to show facilitation as a function of verbal pretraining on the discriminal stimuli. J. Cantor (7), Dysinger (8), Gagné and Baker (10), Goss (13), Price and Lewis (20), and Rossman and Goss (22) did demonstrate facilitation rather conclusively. Examination of the experimental conditions that obtained in these experiments, together with analysis of the theoretical assumptions underlying the ADC hypothesis, suggests at least four conditions which should be met in a test of the hypothesis.

First, the discriminal stimuli must be sufficiently similar that generalization initially occurs from one to the other; at the same time, they must be sufficiently dissimilar that discriminal responses can be learned to them. *Second*, the discriminal stimuli must be of such a nature as not to elicit verbal names or descriptive phrases among the control subjects. If this condition is not met, the experimental subjects will have no advantage with respect to the experimentally introduced names, and may even be handicapped by competing names. It is this condition that seems to make the child the ideal subject for testing the ADC hypothesis, since he is less

likely to use descriptive phrases than is an adult and the selection of appropriate stimuli is thus made easier. *Third,* the responses to be learned in the transfer task should be simple and discrete. If these responses involve mastery of a complex motor coordination, the facilitation contributed by the pretraining may be masked or at least contribute only slightly to the total variation in the learning scores. *Fourth,* if the positive transfer is to be attributed to the effects of the names learned, there should be an adequate control for the "warm-up" or "learning to learn" effects.

A recent experiment to test the ADC hypothesis was conducted by Gerjuoy (11) with elementary school children. Her transfer task consisted of the successive presentations of three stimulus lights, differing from each other only in position on a 20 in. by 20 in. panel, which the child learned to turn off by pushing the appropriate one of three response buttons arranged in a horizontal row. Several experimental groups were given a preliminary task in which they were required to learn nonsense names for the three light positions used in the transfer task. Reasoning that the more similar the names in pretraining, the less effective they should be in the transfer task, Gerjuoy required one experimental group to learn highly similar names while another learned dissimilar names. Some subjects in each of these two groups were required to verbalize the names overtly in the transfer task prior to responding while others were not required to do so. Control subjects were given a preliminary task in which they were simply required to call off the names of colored lights as they appeared on the stimulus panel. Gerjuoy found that all subjects given the name-learning as the preliminary task performed better on the transfer task than those given the color-naming task, but there were no differences in this respect among the experimental groups.

An experiment with preschool children by G. N. Cantor (5) is also relevant to this hypothesis. His transfer task was a simultaneous discrimination problem that involved reinforcement for the selection of one of two toy cars differing from each other with respect to the stimuli pasted on one side. The stimuli were sketches of the faces of two girls with marked resemblance. Relevant stimulus subjects were given pretraining in which they learned to call one of the faces by one name and the other by another name. A second ("attention") group did not learn names for the faces, but had an equal number of trials in which they pointed to various parts of the faces (nose, eyes, etc.), thereby assuring "perceptual" experience with the stimuli. An irrelevant stimulus group learned names for a pair of boys' faces. All three groups were immediately given the transfer task. Cantor found that the group given pretraining with the relevant stimuli learned the transfer task significantly faster than either of the other two groups, and that these latter two groups did not differ significantly from each other.

A subsequent experiment by Norcross and the writer, using Cantor's stimuli, was conducted to determine whether discrimination experience with the relevant stimuli is sufficient to account for the facilitation. The experimental design was like Cantor's except that in place of his "attention" condition, a condition was used under which the subjects were required to learn to say "same" when two identical pictures were presented and to say "different" when the two pictures were different. This procedure assures that the subjects must discriminate between the two stimuli without necessarily giving different names to them. The relevant group learned significantly faster than the irrelevant and discrimination groups, and the latter two did not differ from each other.

This series of experiments, with those conducted on adult subjects, confirms the usefulness of the ADC hypothesis. The experiments by Gerjuoy (11) and others (7, 8, 13, 22) indicate that a facilitation occurs that cannot be attributed to warm-up or a learning set induced by the preliminary task. The experiment by G. N. Cantor (5) indicates that this facilitation cannot be attributed to preliminary visual or perceptual experience with the stimuli, and that by Norcross and the writer suggests further that preliminary discriminatory experience is insufficient to account for the subsequent facilitation. The most important contribution of these experiments lies in the control for perceptual and discriminatory experience. Probably only with young children could one give discrimination experience without at the same time eliciting consistent names or verbal descriptions for the stimuli.

The facilitative effects of stimulus pretraining with names has been shown in the delayed reaction experiment (28). The experimental group was given the usual relevant stimulus pretraining in learning distinctive names for each of a pair of highly similar sets of concentric circles. The control group was given pretraining with these stimuli that involved saying "same" when the two stimuli presented were identical and saying "different" when they were different. The transfer task required the subject to select which of the two stimuli he had seen baited with a reward 25 seconds before. The stimuli were not visible to the subject during the delay period. It was found that subjects given the verbal pretraining selected the correct stimulus significantly more frequently than those without such training, and that this effect was significantly greater for younger than for older subjects. These results are consistent with the ADC hypothesis, as well as with the hypothesis that the names serve as substitutes for the absent stimuli during the delay period.

Assuming tentative verification of the ADC and AEC hypotheses, Shepard (25) made an interesting application of these hypotheses to the classificatory behavior of preschool children. Her transfer task involved five nonsense figures, and the subjects were rewarded for pulling a lever to

three of them and were not rewarded if they pulled the lever to the other two. The relevant stimulus group was given a pretraining task in which they learned one name for the first three stimuli and a different name for the other two. An irrelevant stimulus group was given similar training, but with stimuli not used in the transfer task. The relevant stimulus group made fewer errors on the transfer task than did the irrelevant stimulus group. Shepard accounted for this difference in terms of the ADC and AEC hypotheses. Learning the same name for the stimuli within a class is assumed to increase the amount of generalization that occurred among them. Learning different names for the members of the two different classes is assumed to decrease the amount of generalization between members of the different classes. Both processes thus contribute to facilitation of the relevant stimulus group.

Shepard's application of these two principles to classificatory behavior suggests an explanation for another well-established phenomenon. Kuenne (16) used preschool children to study the transposition phenomenon. As a point of departure, she used Spence's discrimination learning theory (26, 27), which predicts transposition for animals in the two stimulus problem for test stimuli that are similar to the training stimuli, but little or no transposition for test stimuli that are distant from the training stimuli. Kuenne suggested that subjects who are articulate with respect to the stimulus differences may use these verbal responses in such a way that transposition to more distant test stimuli also occurs. She was able to demonstrate, under specified conditions, that the distance of the test stimuli from the training stimuli may be greater, and transposition still occur, for children with higher mental ages (and therefore presumably more articulate with reference to the stimulus relationship) than for children with lower mental ages. A similar finding was reported by Alberts and Ehrenfreund (1) in which the range of transposition was found to be an increasing function of chronological age.

Spiker, Gerjuoy, and Shepard (29) conducted a related experiment in which transposition was studied in the intermediate size problem. They found the amount of transposition to be related to the subject's knowledge and understanding of the concept "middlesized." That is, subjects who knew and understood the concept "middlesized," as determined by a preliminary test, transposed to a considerable degree, but subjects who did not understand the concept failed to transpose to any great degree.

The ADC and AEC hypotheses may be used to analyze the manner in which the verbal concepts affect transposition behavior. Consider the two stimulus problem in which the subject is reinforced for selecting the larger of two stimuli. Suppose that the larger stimulus elicits, by virtue of previous learning, the verbal response "larger one" (or its linguistic equivalent) and the smaller stimulus elicits "smaller one." Suppose, more-

over, that for this subject, the larger and smaller stimuli of any pair of stimuli with discriminable differences in size elicit these distinctive verbal responses. According to the AEC hypothesis, any response (e.g., a choice response) learned specifically to the larger stimulus of a given pair should therefore generalize to the larger stimulus of any other pair, regardless of the distance of the test pair from the training pair. And, because the two stimuli elicit distinctive responses, the generalization between them should also be decreased. Thus, extensive transposition of articulate subjects would be explained by mediated generalization and acquired distinctiveness of cues. A similar derivation can be made for the intermediate size problem.

This account does not, of course, explain how verbal names are learned for stimulus relationships in the first place. Considerable research with young children on such problems seems advisable before speculation is likely to be profitable.

The results of the experiments reported above indicate that research with children can be brought to bear on hypotheses from general behavior theory. Obviously, children are more appropriate as subjects for some problems than for others. There are many factors to consider in selecting problems for research with children. For example, there are some phenomena that are more or less unique, or at least occur in their purest form, in children. The development of the first linguistic behavior can be studied only in very young children. Imitation learning can be studied in children without many of the complicated processes that confound adult imitation. The use of children provides a means of selecting subjects with a level of language ability more under the control of the experimenter than would be true for adults. To arrange for adults to show the same gradual learning as children show in simple problems, the task must often be made many times more difficult at the risk of masking the basic processes. The assumption that the subject is unaware of the purpose of an experiment, the tenability of which is central in numerous experiments, can often be more plausibly maintained if the subjects are unsophisticated children. The development of experimental tasks, for which generalization from previous experiences will be minimal, is undoubtedly easier to accomplish for child than for adult subjects. Certain pragmatic considerations also suggest the more effective use of children in much psychological research. If one is interested in the effects of incentive on performance, dime-store toys for children may be as good or better than dollars for adults. Or, the observable effects of a mild experimental frustration on child behavior may be greater than those of an elaborately arranged "traumatic" experience on adult behavior. Or again, children can usually be more readily convinced than adults that their performance represents success or failure.

Such considerations may result in economy for the experimenter with the ingenuity to make use of them. It may be these factors that many have

in mind when they insist that the child mirrors the basic psychological processes more faithfully than does the adult. Certainly these considerations suggest that, given carefully developed apparatus, procedure, and research methodology, the child may be one of the most fruitful subjects for psychological experimentation, not only in the investigation of practical problems as in the past, but also in the testing of hypotheses that stem from general psychological theory.

References

1. Alberts, E., & Ehrenfreund, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, 41, 30-38.
2. Arnoult, M. D. Transfer of predifferentiation training in simple and multiple shape discrimination. *J. exp. Psychol.*, 1953, 45, 401-409.
3. Birge, J. S. The role of verbal responses in transfer. Unpublished doctor's dissertation, Yale Univer., 1941.
4. Bugelski, B. R., & Scharlock, D. P. An experimental demonstration of unconscious mediated association. *J. exp. Psychol.*, 1952, 44, 334-338.
5. Cantor, G. N. The effects of three types of pretraining on discrimination learning in preschool children. *J. exp. Psychol.*, 1955, 49, 339-342.
6. Cantor, Joan. The effects of pretraining on the performance of a simplified Mashburn task. Unpublished master's thesis, State Univer. of Iowa, 1953.
7. Cantor, Joan. Amount of pretraining as a factor in stimulus predifferentiation. *J. exp. Psychol.*, 1955, 50, 180-184.
8. Dysinger, D. An investigation of stimulus predifferentiation in a choice discrimination problem. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
9. Eisman, Bernice S. Attitude formation: The development of a color-preference response through mediated generalization. *J. abnorm. soc. Psychol.*, 1955, 50, 321-326.
10. Gagné, R. M., & Baker, K. E. Stimulus predifferentiation as a factor in the transfer of training. *J. exp. Psychol.*, 1950, 40, 439-451.
11. Gerjuoy, I. R. Discrimination learning as a function of the similarity of the stimulus names. Unpublished doctor's dissertation, State Univer. of Iowa, 1953.

12. GERJUOY, H., & BECHTOLDT, H. P. Generalization effects in successive discrimination learning as measured by response similarity. Department of Psychology, State Univer. of Iowa, 1955.

13. GOSS, A. E. Transfer as a function of the type and amount of preliminary experience with task stimuli. *J. exp. Psychol.*, 1953, 46, 419-428.

14. HULL, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.

15. JEFFREY, W. E. The effects of verbal and non-verbal responses in mediating an instrumental act. *J. exp. Psychol.*, 1953, 45, 327-333.

16. KUENNE, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, 36, 471-490.

17. MCALLISTER, D. E. The effects of various kinds of relevant verbal pretraining on subsequent motor performance. *J. exp. Psychol.*, 1953, 46, 329-336.

18. MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, 43, 155-178.

19. MURDOCK, B. B., JR. Effects of failure and retroactive inhibition on mediated generalization. *J. exp. Psychol.*, 1952, 44, 156-164.

20. PRICE, H. G., & LEWIS, D. Increased pronouncing behavior as a factor in serial learning. *J. exp. Psychol.*, 1954, 47, 95-100.

21. ROBINSON, J. S. The effect of learning verbal labels for stimuli on their later discrimination. *J. exp. Psychol.*, 1955, 49, 112-114.

22. ROSSMAN, I. L., & GOSS, A. E. The acquired distinctiveness of cues: The role of discriminative verbal responses in facilitating the acquisition of discriminative motor responses. *J. exp. Psychol.*, 1951, 42, 173-182.

23. RUSSELL, W. A., & STORMS, L. H. Implicit verbal chaining in paired-associates learning. *J. exp. Psychol.*, 1955, 49, 287-293.

24. SHEPARD, W. O. Mediated generalization with high interstimulus similarity. Unpublished master's thesis, State Univer. of Iowa, 1953.

25. SHEPARD, W. O. The effects of verbal pretraining on discrimination learning in preschool children. Unpublished doctor's dissertation, State Univer. of Iowa, 1954.

26. SPENCE, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, 43, 427-449.

27. SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

28. SPIKER, C. C. Stimulus pretraining and subsequent performance in the delayed reaction experiment. *J. exp. Psychol.*, in press.

29. SPIKER, C. C., GERJUOY, I. R., & SHEPARD, W. O. Transposition by children with the intermediate size problem. *J. comp. physiol. Psychol.*, in press.

PSYCHOLOGICAL DEVELOPMENT IN MIDWEST

HERBERT F. WRIGHT
University of Kansas

Midwest is a small town in a midwestern farm region of the United States. Some 100 of its approximately 700 inhabitants are children under 12 years of age. Although the supply lines of American culture and technology run through it, Midwest nonetheless has essential independence and completeness (2, pp. 20-32), with the result that it holds the lives of these children almost entirely within its borders. One confronts an "intelligible field of study" (4), therefore, in asking the question raised here: What age related changes occur in the behavior and psychological living conditions of Midwest children? What is it like to grow up in Midwest?

The Midwest Field Station has been working for some time with this and related problems in psychological ecology.[1] Its special task has been to describe along as many dimensions as can be accommodated by available procedures and concepts the psychological habitat and naturally occurring behavior of Midwest children. This it has aimed to do with something of the outlook and aspirations of the weather outpost and of biological and astronomical field stations which have reaped from continuing records of naturally occurring phenomena substantial benefits in positive knowledge and basic theory. Methods of the research and results of studies to date have been reported elsewhere (see especially 2, 3, 11, 12). I will try here only to review some of these methods and to bring together some of their

[1] The work reported here has been directed by Roger G. Barker and the author, and supported in part by the Institute of Mental Health of the National Institutes of Health, U.S. Public Health Service, and the University of Kansas. Other direct contributors to the studies represented include William A. Koppe, Jack Nall, and Phil Schoggen.

results that answer only to the one problem of change with age in behavior and its conditions in Midwest.

Developmental Change and Behavior Settings

Behavior Settings as Community Units

Procedures to be reviewed below indicate that of their person hours in one recent year Midwest citizens of all ages spent

18,525	in Clifford's Drug Store,
7,600	in the Second Grade Classroom,
7,300	in Hooker's Tavern,
3,750	in the Methodist Regular Worship Service,
1,000	in the Midwest Cemetery,
540	in the City Library, and
344	in the Brownie's Regular Meeting

of this community. Here are plainly visible Midwest centers of human activity. We have called such action areas *behavior settings,* and we have used them to describe some aspects of psychological development in Midwest.

A behavior setting can be defined as a stable part of the physical and social milieu of a community together with an attached standing pattern of human behavior. Always the milieu part has its own distinctive characteristics of time, place, things, and personnel that accommodate the behavior pattern. Always, too, this pattern is extraindividual and invariant in that it stays essentially the same while different persons come and go.

Take as a single example the Methodist Regular Worship Service. Its time is 11 to 12 a.m., every Sunday, and its place the Methodist Sanctuary. Its things include hymn books, collection plates, and a pulpit. Ushers, congregation members, the choir, and a minister are it personnel. These elements form the milieu part. The behavior pattern includes ushering, sitting, standing for prayer, preaching and listening, gathering in the vestibule, conversing, and shaking hands, all of which change little Sunday after Sunday, year after year, no matter who goes to church.

Behavior settings are common empirical facts of such prominence that independent observers identify them and most of their features and functions with near unanimity. Yet each contains forces that coerce individual behavior, as abrupt changes in the action streams of individuals when they leave one Midwest setting and enter another regularly imply. Collectively, moreover, a community's settings have a kind of omnipotence in that virtually all things done by its citizens occur within their limits. For these and other such reasons (3, pp. 45-57) we have found in behavior settings appropriate analytical units for research on some aspects of the natural history of human behavior in Midwest.

A Behavior Setting Survey

All of Midwest's behavior settings, from New Year's Holiday Dinner to Santa Claus at the Court House, occur at least once in the course of a year. The station has accordingly made a behavior setting survey of Midwest for a town year: July 1, 1951, to June 30, 1952. There were identified during this survey year 2,030 Midwest settings, of which 1,445 occurred in homes and 585 in more public areas. Special criteria were used to verify the existence of different settings as discrete units of the same magnitude via tests of their mutual independence (3, pp. 62-66).

Chief consideration will be given here to the 585 public units or *community behavior settings* in which, according to our data, Midwest children spent approximately one-third of their waking time. These were grouped into 107 *varieties*, among which there were *School Classes, Indoor Entertainments, Restaurants and Taverns, Barber and Beautician Shops, Sunday School Class Meetings, Circuses and Carnivals, Organizational Meetings With Refreshments, Organizational Meetings Without Refreshments, Parades, Funeral Services, Food and Rummage Sales, School Field Trips*, and *Fire Drills*. Seventy of the 107 varieties and approximately 205 member settings were found to occur in the town during an average week day.

TABLE 1

AGE GROUPS OF THE MIDWEST POPULATION

	Number of Persons		
	Males	*Females*	*Total*
INF: Infant (birth to 1 year, 11 months)	15	10	25
PRE: Preschool (2 years to 5 years, 11 months)	18	12	30
YNG: Younger School (6 years to 8 years, 11 months)	14	15	29
OLD: Older School (9 years to 11 years, 11 months)	14	21	35
ADO: Adolescent (12 years to 17 years, 11 months)	25	25	50
ADU: Adults (18 years to 64 years, 11 months)	172	204	376
AGD: Aged (65 years and over)	82	80	162
All ages	340	367	707

NOTE.—The different age groups are identified in subsequent figures by the abbreviations used here.

The procedures of the behavior setting survey can be used to describe aspects of the behavior and life situations of adolescents and adults as well as of children. Therefore, to get information on behavioral change throughout the life span in Midwest, the station has applied these procedures to all of the seven age groups of Table 1, with results now to be sampled.

Time in Settings and Age

We can begin with a plain question. What changes occurred with age in the time Midwesterners spent in different behavior settings?

The answering data are based on estimates of *occupancy time*, anticipated by the itemization above of person hours in selected settings of Midwest. This measure was obtained by multiplying the number of occurrences of a setting during the year (its frequency, f) by the number of Midwest citizens who entered it during each occurrence (its population, p) by the time in hours these persons spent in the setting when they entered it (its duration, d). The product fpd gave the person-hours-per-year or the occupancy time of the setting. Separate occupancy measures were computed for each age group and other subgroups of the population. Also, the sum of the time spent in all community settings was then determined for every such group. Finally, the percentage contributed to this sum by each setting was computed as a group *occupancy index* for that settings unit.

TABLE 2

OCCUPANCY INDEXES OF THE TEN COMMUNITY BEHAVIOR SETTINGS IN WHICH MIDWEST PRESCHOOL CHILDREN AND MIDWEST ADULTS SPENT THE MOST TIME

OCCUPANCY INDEX:	PER CENT OF	
22,420	593,994	
Trafficways 21.4	8.1	..... Trafficways
Culver Shoe Repair 6.7	2.9	..... Telephone Office
Midwest Theater 6.2	2.6	..... Garnett's Grocery
Mrs. H Baby Care 4.5	2.6	..... Kanes Grocery
Methodist Kindergarten 3.3	2.5	..... Gwynn Cafe
Mrs. D Baby Care 3.3	2.5	..... Bank of Midwest
Presbyterian Kindergarten 2.0	2.4	..... Wherry Door Company
Clifford's Drug Store 1.7	2.4	..... Midwest Theater
Pearl Cafe 1.6	2.4	..... Clifford's Drug Store
Saturday Night Dance 1.3	2.1	..... Midwest Hardware
Preschool Children		*Adults*

NOTE.—Trafficways are streets and sidewalks. The Culver Shoe Repair housed the Culver family whose eleven (11) children had many friends. The kindergartens are Sunday-School classes. Children often tag along to dances in Midwest.

Table 2 lists, first, the 10 community settings in which Midwest preschool children spent the most time and, second, the 10 in which Midwest adults spent the most time. Each setting is named with its occupancy index. Some impression as to what growing up brings in Midwest may be gained by comparison of the two lists in this table.

Our occupancy figures indicate that Midwesterners spent 5,130,000 hours (waking and sleeping) in family settings and 1,030,658 hours in community settings. Data on the average number of hours per day all residents occupied these Midwest living areas are summarized for seven age levels in Figure 1. Each average is for all weekdays, holidays, and Sundays of the entire year.

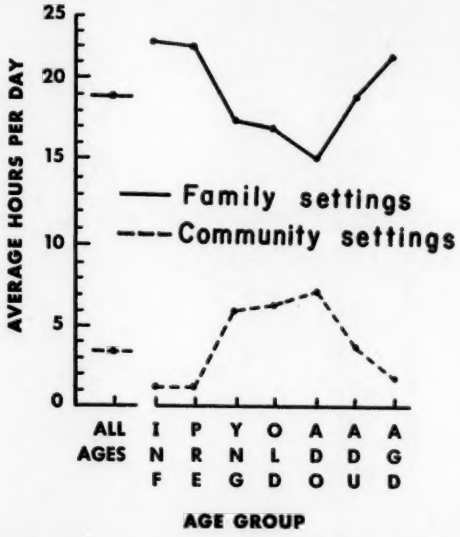

FIGURE 1—Average hours per day spent in family and community behavior settings by the total population of Midwest and by different age groups.

Figure 1 shows that family settings surpassed community settings about 5 to 1 in competition for the time of Midwest citizens. But it also shows dramatic change with age in this competition. Temporal preeminence of the family declined greatly from infancy and early childhood through adolescence, but then rose above the maximum level for school children in adulthood, and approached the level for infants in old age. One might easily be surprised at first to find older children and adolescents spending, as Figure 1 indicates, much more time than their elders away from home. Are not the years of later childhood and youth family years? The answer, of course, is that they are school years as well. It is clear in any case that growing up in Midwest through adolescence was considerably a matter of growing into public behavior areas of the community.

Number of Settings and Age

What changes occurred, then, as Midwesterners grew older in the number of different community settings they entered? How did freedom to roam over the town vary with age?

This question the station has answered for each of several Midwest population subgroups in two steps. The *territorial range* of the group, defined as the number of community settings inhabited during the survey year by one or more group members, was first determined. This measure was then used to calculate as a *territorial index* the percentage of all community settings (585) entered by individuals of each group. The territorial indexes of all of the different age groups are presented in Figure 2.

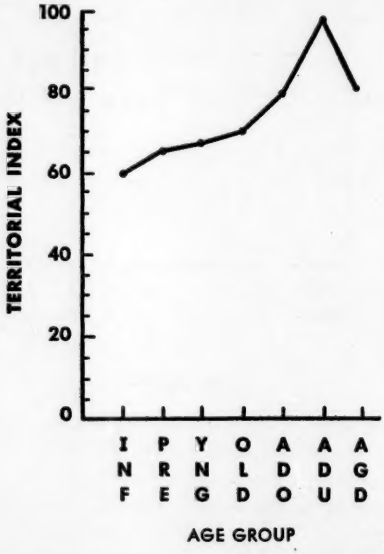

FIGURE 2—Percentage of all community behavior settings inhabited by members of each group.

A Midwest four-year-old once informed a station observer who had expressed astonishment at seeing him on his own far from home: "I go over the whole world." Figure 2 indicates that something like this was true for Midwest children of every age. Even for infants the territorial index was 60; and this indicator of free space rose regularly and without any sign of sudden or radical shifts of scene to 79 for adolescents.

Because territory to roam in did grow with its recipients in Midwest, children of each age level had ahead of them behavior settings beyond the occupied free area. Getting to be six when one was three, nine when one was six, and so on up promised a larger field of action. We do not know whether or how much cities differ here from Midwest. Data for comparison are available, though, from a smaller community, the Lawton home and school for crippled children (8). There, to begin with, the total number of settings was but 59. The oldest residents, 14-year-olds, entered nearly all of these settings, as shown by a territorial index for them of 95. But so it was precisely for the youngest residents, who were six-year-olds! Growing up in Lawton had little to offer. Growing up in Midwest both offered and brought continuous expansion in behavior settings.

Penetration into Settings and Age

The finding that Midwest children occupied an increasingly larger number of community behavior settings as they grew older tells nothing of how far they got into these settings and so leaves open the question as to change with age in level of involvement and responsibility in Midwest's public behavior areas.

We have rated the maximum *depth of penetration* by members of different age groups into community settings of the town on a scale that extends through six penetration zones. These zones of penetration, listed in ascending order according to their centrality or depth, and identified in terms of characteristic participation roles, with each exemplified by a single participant, are as follows: (1) *onlooker* (loafer in the Post Office); (2) *audience participant or invited guest* (spectator at a ball game); (3) *member or customer* (scout at a scout meeting); (4) *performer or active functionary* (actor in a play); (5) *joint leader* (co-proprietor of a store); (6) *single leader* (teacher of a class). The maximum depth of penetration of each age group was defined as the deepest zone reached by any member of that group during the survey year.

Figure 3 represents the distribution of the maximal depths of penetration into community settings at each of the seven age levels. It reveals an increase from infancy to adulthood and a retrogression in old age to about the adolescent level in depth of settings penetration. Again, as in expansion of territorial range, the gains made through childhood and adolescence were regular and gradual. Participations in the onlooker and guest zones (zones 1 and 2) went down while those in the member and performer zones (zones 3 and 4) went up at a slow and remarkably constant rate.

But Figure 3 points to another more conspicuous fact. It is that to stand much chance of being a leader or co-leader of any kind in Midwest you had to be an adult. Children of every age and adolescents put together penetrated to the leader zones (zones 5 and 6) in only 8 per cent of all community

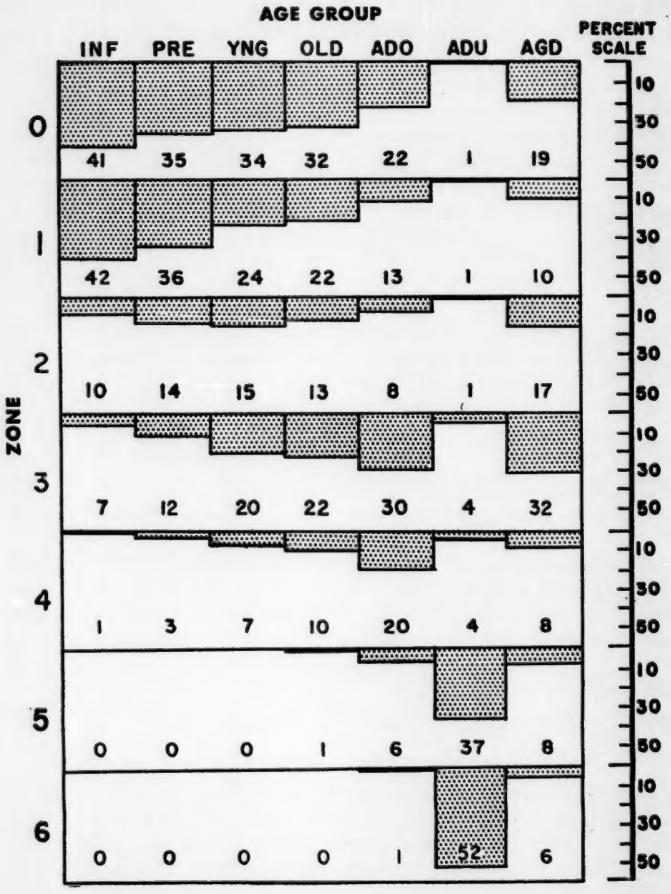

FIGURE 3—Penetration of community settings by Midwest age groups. Percentage of all community settings which age group members entered to different maximal depths of penetration. Zone 0 represents the community settings not entered by any member of an age group.

settings. Adults, by contrast, occupied these power centers in 89 per cent of all community settings.

So the freedom to roam of Midwest children was not balanced by freedom to direct. The bars in zones 5 and 6 show that, as a matter of fact, adults filled most of the leadership openings Midwest had to offer. This meant that grownups were nearly always on hand to look after children. But it meant also that children were not often left to look after themselves, with power to manage their own affairs in their own dens, clubs, gangs, or other child maintained behavior settings. Current theories are too equivocal to permit anyone to say categorically whether this was good or bad for psychological development in Midwest. It can be said that Midwest children got into less trouble than some children of cities who are reported to be "on the loose" and without adequate "supervision." Yet it is possible that they could have been better off with less adult surveillance and more independence to form and lead child groups in settings free of adults.

Figure 3 shows in any case that children in the public parts of Midwest were by no means only bystanders without status or responsibility. Older school children were at least members or customers (zone 3) in 22 per cent of all community settings. So were younger school children in 20 per cent, and even preschool children in 12 per cent. Also, 55 per cent of the settings children and adolescents penetrated to the performer zone (zone 4) were *County Teachers Institute, County Welfare Office, Women's Club I,* and other such fully "adult" centers. Children were indeed welcomed and even pressed into these settings as entertainers, clerks, office helpers, assistant mechanics, janitors, and even type setters. Our observations suggest, moreover, that for this reason the self esteem of Midwest children was high even though their power was low in most community settings.

Segregation in Settings and Age

There occurred relatively little segregation of Midwest inhabitants into narrow age, sex, social class, or race groups in community settings. But what there was of this changed some with age.

A quite regular decrease occurred from 90 in infancy to 53 in adulthood, with a reversal back to 67 in old age, in the per cent of community settings not limited in any degree (e.g., Pearl Cafe) as to age group.

Throughout childhood and adolescence, without a jog in the curve, 92 per cent of the settings entered by Midwesterners of this range were inhabited by both sexes, with this figure going down to 90 in adolescence, on down to 81 in adulthood, and back up to 84 in old age.

Ninety-seven to 98 per cent of the settings entered by children of each age group and by adolescents were inhabited by all three of the town's social groups (which correspond closely to the Upper Middle, Lower Middle, and Upper Lower categories of Warner (10)). The corresponding per cents for adults and the aged respectively were 93 and 94.

During childhood and adolescence 60 to 66 per cent of all community settings were inhabited by Negro residents (of whom there were 27) and white residents, while 54 per cent of these settings were nonsegregated as to race in both adulthood and old age.

It is clear that barriers between age groups, the sexes, the three social groups, and the two races were in no case very high in Midwest, with the exception perhaps of the race limitation. But it also is clear that age related forces tended to raise all such barriers.

Standing Behavior Patterns and Age

Differences were found from age to age in characteristics of the standing behavior patterns of community settings within the territorial ranges of the seven age groups. Demonstration of these differences rested upon ratings with suitable scales of the behavior pattern anchored within every setting.

The salient molar components of *action patterns* in community settings *entered* by infants were social contact (sociability, fellowship), recreation (pleasurable activity, games, entertainment), nutrition (dining in public places), religion (worshipping, religious study and teaching), and aesthetics (beautifying surroundings or self). A gradual shift in the relative prominence of different action patterns occurred through childhood to adulthood to old age, where the most prominent patterns were social contact, earning a living (hiring out for wages), recreation and government (creating, interpreting, and enforcing laws and regulations).

The molecular components or behavior mechanisms outstanding in community settings entered by infants were talking (verbalization of all kinds, including singing and yelling), affective behavior (overtly expressed emotionality), and thinking (decision making and problem solving). Again, a process of gradual change took place until, in old age, the most prominent mechanisms in the inhabited settings were thinking, gross motor activity, and manipulation (use of the hands in prehension).

Both of these courses of change, the one in goal directed action, and the other in act mechanisms, point to a movement from relatively little involvement, during the earliest years, to relatively great involvement, in old age, with necessities of the real world.

Developmental Change in Behavior Episodes

Episodes as Individual Behavior Units

Psychological development in Midwest, while it proceeds in behavior settings as community units, necessarily involves all of the time units of behavior and situation in the lives of individual persons. Here also the Midwest behavioral panorama includes readily observed targets of description, which we have called *behavior episodes*. These have been used as

standard ecological units to take into account the behavior and psychological habitat of individual Midwest children.

To collect episodes for study we have employed the *specimen record,* a detailed, sequential, narrative account by skilled observers of an individual child's behavior through a more or less extended time. At its longest a record of this kind spans a day—and some 300 pages. Here is a brief sequence from a specimen day record on Mary Ennis, an eight-year-old girl. It occurs in the Ennis home, to which Mary has just returned from school with a box of paints and other belongings. Mary stands in her bedroom, from which one door leads to a dining room, and another to a bathroom. Her mother is seated in the dining room, and the observer stands across the bedroom from Mary.

Painting Lips

Mary took her box of paints and went directly and purposefully into the bathroom.

She turned on the water.

A mirror was behind the door and because the door was open a little, I could see through the crack that Mary was standing at the mirror.

Mary said, "I'm trying to do something," and giggled a very mischievous giggle.

Her mother evidently recognized the quality of the giggle because she came into the bedroom to see what was going on.

We both saw that Mary was painting her lips with the paint brush and the water color paint.

Washing Paint Off Face

Mrs. Ennis said, "Oh, you had better not do that. That will make your lips chapped." She added for good measure, "I don't think it will come off," and then left the room.

Immediately, almost, Mary stepped from the mirror to the sink. She started to run the water and wash off the paint.

Challenging Mother

She said, "What did you say? Did you say it wouldn't come off?" This was a challenge to the mother, which Mary repeated with a good deal of emphasis, saying, "Mommie, you said it wouldn't come off and it did."

The mother answered, "I didn't think it would." She wasn't arguing, just explaining.

Mary worked still at the sink. She turned on the water full force and let it run for a full minute while she washed off the brush and fussed around with the paints.

Then she turned the water off and came back into her bedroom.

Readers will have noted that this sequence is marked off by side brackets into three parts, entitled: *Painting Lips, Washing Paint Off Face,* and *Challenging Mother.* These we take to be episodes. Other episodes of our records include *Working Problem at Board, Riding Father's Foot, Making Rock Garden, Getting Splinter Out of Hand, "Painting" Porch with Mud,* to which I will return later, *Telling Off Friend, Sweeping Floor,* and *Getting Dressed.*

Every episode contains an action and a particular situation that correspond respectively to the standing behavior pattern and the milieu context of a behavior setting. The action is always a segment of behavior with a constant direction, and the situation is the coexisting set of conditions in the person and the environment which interact to determine the action. Take *Washing Paint Off Face,* of Mary's day. This title identifies the action and indicates its direction. The situation includes on the person side the need state implied by Mary's going on her own to the sink. It indicates on the environment side the cautioning mother, the warning she expressed, the familiar rooms, the mirror, the paints, the sink, and the water as carriers of psychological effects. There are as parts of the situation, in short, all of the conditions that made a difference to the behavior of this child at the time, as these are represented in the record. From this much it will be understood further that an episode is a *molar* behavior phenomenon as defined by Tolman (9) and Lewin (6, 7).

Conceptual and operational criteria of episodes have been developed in some detail (3, pp. 225-270). These criteria have been tested with resulting evidence that the process of discriminating episodes in specimen records is reliable (3, pp. 270-273).

Recorded episodes can be studied with sole regard for their action and situation content. But they also can be studied from a quite different point of view. To divide a specimen record into episodes is to find parts of a whole behavior sequence. Inquiry can then be made into relations of part to whole and part to part within the sequence without any reference to what the parts contain. This, in the terminology we have used, is to ask about the structure or anatomy of behavior. We have analyzed the episodes in records on Midwest children from the standpoint of both their action and situation content and their structural properties, with consequent findings in each case on change in behavior and habitat with age. I will try, therefore, to represent here each of the two types of analysis, beginning with the anatomy problem.

Behavior Structure

The operation of episoding opens this problem to study by indicating on the margin of a specimen record some basic whole to part and part to part relations. These can be reviewed quickly.

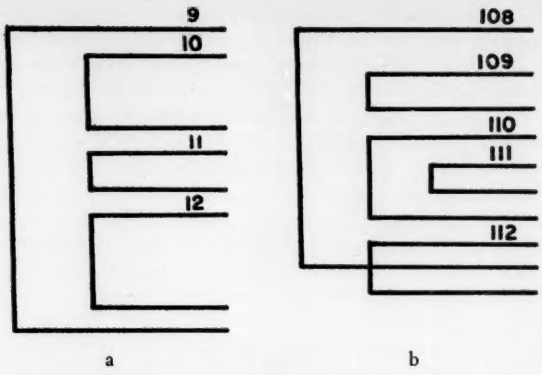

FIGURE 4—a: An episode linkage from a day of a Midwest two-year-old.
b: An episode linkage from a day of a Midwest ten-year old.

First, different episodes can occur at one time, in which case they are said to *overlap*. Overlapping is graphed by parallel placement of episode brackets. The brackets that bound *Washing Paint Off Face* and *Challenging Mother* of Mary's day, for example, are so placed to show overlap such that the whole of the shorter unit occurred while a part of the longer one was in progress. As many as four episodes, though rarely more than three, have been found in progress at one time in our records. Three were judged to overlap, for example, while Wilma sat eating toast, petting her dog, and asking if she could go outdoors to play. Where, as in the case of Mary's *Painting Lips*, a single episode occurs alone, we say it is *isolated*.

Any combination of episodes which overlap to form an interwoven continuum we have called a *linkage*. The sequence from Mary's day extends through two linkages of which the first has only one *link*. Different episodes of a linkage often overlap sequentially with a single continuously ongoing unit. See Figure 4a. It represents only by episode brackets a part of a day record on two-year-old Mary Chaco. At 7:31 A.M., the ninth unit of the day began. Then, while it continued, Mary became involved successively in episodes 10, 11, and 12. A linkage can be much longer and more complex. Figure 4b graphs a just moderately intricate one, presenting unit to unit relations which readers themselves may note, from a day record on 10-year-old Douglas Crawford.

A sketch of our main findings on behavior structure will provide a basis for description of differences age made in this area.

The number of episodes per day varied from approximately 400 to about 1,300 through a median of 850 (after correction for slight inequalities

in length of day). These episodes ranged in duration from a few seconds to 121 minutes through a median of 45 seconds, with 70 per cent shorter than two minutes. Linkages averaged about 400 per day. These sequences of structurally interrelated units included as many as 81 episodes, with the median number 2.3; and their average duration was about two minutes. Fifty to 75 per cent of the episodes per record overlapped with at least one other. Only 10 per cent of these concurrent units, however, overlapped with more than two others. Roughly 11 per cent of the episodes per record were left incomplete, after nonattainment, failure, frustration, or interruption from without. Two per cent or less in each day were broken off and then resumed later. When this happened, the number of interpolated episodes varied greatly, but made up no more than 10 per cent of the units in any record.

TABLE 3

RANK ORDER CORRELATIONS BETWEEN AGE AND VARIABLES OF BEHAVIOR STRUCTURE

Area	*Variable*	*tau*	*p*
Unit differentiation			
	Number of episodes	−.61	.01
	Number of linkages	−.67	.01
	Episodes per linkage	.73	.01
Unit duration			
	Minutes per episode	.61	.01
	Minutes per linkage	.67	.01
Position factors			
	Episodes overlapping	.61	.01
	Episodes overlapping with five or more others	.50	.03
	Episodes between interrruptions and resumptions	.42	.07
Episode termination			
	Episodes left incomplete	−.74	.01
	Episodes ending with another ongoing	.64	.01

NOTE.—In this table and hereinafter the coefficient is *tau* and the *p* test is two-tailed.

The rank order correlation coefficients of Table 3 sum up our findings on change in behavior structure with age. These coefficients show or at least suggest that psychological correlates of increasing age brought about for the older children as compared with the younger: fewer but longer episodes; fewer but longer and also more complex linkages of episodes; more frequent occurrences of different episodes at the same time; fewer incomplete episodes; and more episodes between the interruptions and the resumptions of discontinuous episodes—with the last implying, perhaps, greater ability to sustain need tensions.

These data are consistent with the expectation of greater complexity in the behavior of older children and with commonly mentioned marks of maturity. So it appears especially from the findings that the younger children tended more than the older to do things sequentially, one at a time, to move often from one action to another, and to leave undone what they started, while the older children tended more than the younger to get involved in more than one thing at a time, to engage in lasting actions, and to complete what they set out to do.

The results here open speculation about their meaning for the psychological ecology of Midwest as compared with other communities. Comparable results from children on isolated farms, say, or from children of the Loop in Chicago could have been quite different. Complex overlapping of episodes or extended interlinking of episodes, for example, may occur less often on the farm, where behavior settings are scarcer, reputably more simple, and less closely jammed together. Both of these complications, on the other hand, may occur more often in the Chicago Loop. Such differences might conceivably be used, moreover, as partial indexes of more profound dissimilarities in the tempo, the complexity, and perhaps the stresses and strains of community life. This can be only guessing now. But it has seemed to us that anything as intelligibly sensitive to age differences as behavior structure appears to be must also be sensitive in some ways to culture and community differences.

There remains a larger outcome of this exploring which may have potentialities for research on general principles of psychological development. It is simply demonstration that the psychological continuum has fundamental properties which can be described reliably in terms of behavior structure. How many things does a child do in a day? How are these things related as continuum units? We do not know that such questions have been asked earlier. But our present point is only that these questions evidently can be answered in a reliable way. To be added here to the earlier mentioned evidence that specimen records can be episoded reliably there is evidence that application of the present descriptive categories also is reliable (2. pp. 274-285). That the data have internal consistency is pertinent as well in this connection. So finally are indications of agreement between the data and theories resting on independent and quite different information. This holds especially for what we have taken to be affinities (discussed in 2, pp. 300-302) between our structure findings and developmental theories of Lewin (7), Kounin (5), and Baldwin (1).

Psychosocial Habitat and Behavior

Episodes are far from being the neutral elements which their structural analysis as parts of the behavior continuum might suggest. Two ways in which this is true are important here. First, psychological development oc-

curs in and through episodes. Second, episodes often are arenas for face to face social interaction.

We have used episodes in specimen records on both of these counts in a study of the psychosocial habitat and behavior of Midwest children. Our data indicate that 60 to 80 per cent of these action units involve the children in relationships with mothers, fathers, teachers, peer friends, neighbors, pets, and others of Midwest. What are the properties of these associates as carriers of psychological effects? This is chiefly a question of how others behave in relation to the children. But how, also, do the children behave in relation to others? And what kinds or patterns of interaction per se occur as between the children and their associates?

We have applied to recorded episodes of Midwest children a set of categories dealing with elemental aspects of social action and situation to get some answers to these questions. An example may be helpful. The mother of Margaret Reid, a Midwest four-year-old, orders Margaret to stop "painting" a porch pillar of her house with mud. Margaret says, "I want to paint." This interaction continues and gets increasingly complicated in the course of a troubled episode lasting six minutes. The following are the main questions asked about this episode, as about others, by means of the categories at hand.

As between the mother and Margaret, what is the ratio of **power**? In other words, as Margaret saw it, who of the two has the greater potential ability to change the behavior of the other? A rating is made on an 11-point scale, with equal power at its center.

What degree of "used power" or **social pressure** is the mother actually bringing to bear on Margaret, and vice versa. A rating is made here on a 5-point scale.

What **mode of action** does each of these interacting persons exhibit? *Dominance,* defined as authoritarian pressure? *Aggression,* defined as direct physical or verbal attack? *Submission,* to either dominance or aggression? *Resistance,* against either of these? *Appeal,* for help, a gift, or protection? *Nurturance,* in the sense of acting to benefit the other person by extending help, a gift, or protection? *Avoidance,* in the sense of hanging or drawing back or withdrawing from, the other person? Presence or absence is judged on every action mode of these seven. Also, the same judgment is made on *compliance,* defined as doing what the other person wants as a response to appeal rather than to authoritarian pressure.

What degree of momentary positive or negative *affection* does Margaret show for her mother, and the mother for Margaret? The same is asked about *evaluation* in the form of approval or disapproval by each of the other's conduct. And what is the *mood* of each in terms of degree of happiness or unhappiness? Each of these three **attributes of action**—affection, evaluation, and mood—is rated on an 11-point scale, with neutral at its center.

Some secondary questions also are raised.

In what **behavior setting** does the episode occur? The answer here: Home Outdoors.

Is the **action circuit** *open* or *closed*? Is there action, in other words, by only one of the associated persons, or by both? Margaret responded to the mother, so the answer here is *closed.*

What level of **associate complexity** is there on the associate side, *simple* or *compound*? If simple, the child is acting in relation to one person; if compound, in relation to two of more persons, as when a pupil reports to a class. In the present case, the mother made up the whole of Margaret's social behavior object, which means that the answer here is *simple*.

Who took the **initiative** in the transaction? The associate or the child? Here, obviously, the associate.

These and ancillary categories, which include some to identify both subject and associate, plus others to describe the target unit in terms of its structural properties, we have applied wherever each was applicable to approximately 10,000 episodes of specimen day records on four boys and four girls differing in age from two years through ten. Tests made on representative samples of these episodes have shown varying but generally satisfactory agreement between independent analysts in application of the categories. As in reporting the results on behavior structure, I will start here with a partial summary of overall findings, and then present some data on age relationships.

Sixty to 80 per cent of the episodes in every day involved the children in relations with an associate. A complete social interaction, such that the child and the other person each acted in relation to the other, occurred in approximately four-fifths of these episodes. Associates of the children usually were single individuals rather than groupings of two or more individuals.

Adults far exceeded other children as associates. The most nearly omnipresent among these adults were mothers, who figured in about three times as many episodes as fathers. Teachers were second only to mothers in the number of episodes they entered. As these findings suggest, greater rated power of associates was quite generally characteristic of the children's social situations. Adults, especially, greatly exceeded the children in power. The children nonetheless initiated approximately half of their interactions with individuals of every associate grouping.

The data on modes of action are of special interest here. Figure 5 represents the frequency of each mode in action by, and in action of the children toward, both adult associates and child associates. All of the different modes are identified by prototypic verbal expressions.

Two complementary relationships appear in the mode findings for episodes with adult associates. First, a *control* relation is indicated by the evidence that adults were often dominating while, at the same time, the children were often submissive. Second, a *dependency* relation is indicated by the evidence that children often appealed to adults while adults were often nurturant. These two relationships, the one between "bossing" and giving in, and the other between asking for benefit and getting it, can be brought together in the conclusion that the social regimen of the children with adults was a kind of *benevolent autocracy*. This regimen had, further, the common authoritarian feature of power centralization. As data which cannot be added here show, the power was centered preeminently in mothers and teachers; the autocracy was matriarchal. Readers may find, however,

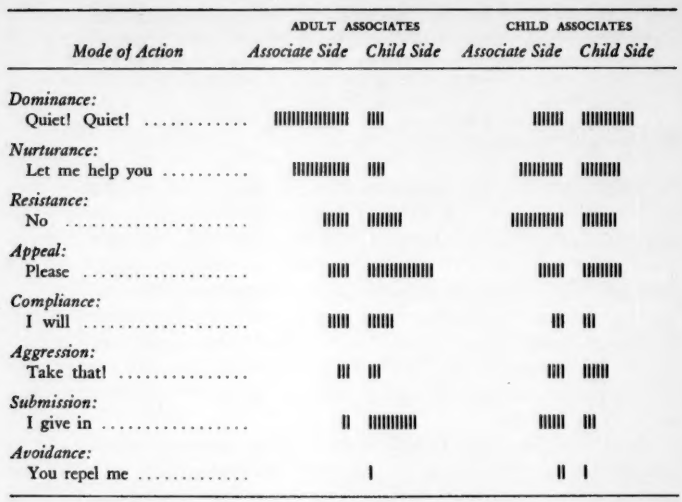

FIGURE 5—Modes of action by Midwest children and their adult and child associates. Each row of symbols represents the number of ocurrences of a mode in every 50 episodes. For example, the row of symbols on the child side, under adult associates and opposite dominance, shows that in four episodes of every 50 dominance occurred in behavior of the children with adult associates.

as we have concluded by studying Figure 5 and related data, that this adult with child situation was not characteristically subjugative or bridling (note high child resistance) or hostile (note high adult nurturance, low adult aggression) or rejective (note zero adult avoidance) or conflictive (note low counter dominance on the child side, low aggression on both sides).

Figure 5 shows that the social pattern of children with children was not featured by an authoritarian control relation. Thus, although child associates were often dominative, the children were not at all equally submissive. The subjects tended more to be resisting, and they were even more dominative than their child associates. The possibility of a transfer to weaker individuals of impulses to oppose adult authority is suggested by relatively more frequent dominance, aggression, and resistance by the children against children than by the children against adults.

Data from the *mood, affection,* and *evaluation* categories indicate that the social situations of the children were preponderantly pleasant, friendly, and approving, in line with the evidence of frequent nurturance from both

adults and other children. Also, findings on *kind of interplay* indicate that interactions of the children with others were harmonious much more often than disharmonious, although outright conflict, discord short of conflict, and unfriendly rivalry did occur in these interactions with considerable frequency. One apparent reason for this is that, as the data on relative power and the indications of authoritarian control by adults lead one to anticipate, *social pressure* by others on the children greatly surpassed in frequency and strength social pressure by the children on others.

Table 4 presents coefficients which indicate a number of changes with increase in age of the children in these various aspects of social behavior and situation. The more prominent of these changes follow.

TABLE 4

RANK ORDER CORRELATIONS BETWEEN AGE AND PSYCHOSOCIAL VARIABLES

Area	*Variable*	*tau*	*p*
Associate identity			
	Episodes with adults	−.79	.01
	Episodes with mothers	−.71	.05
	Episodes with fathers	−.64	.05
	Episodes with children	.64	.05
	Episodes with compound associates	.82	.01
Social habitat			
	Average power of all associates	−.93	.01
	Average power of child associates	−.71	.05
	Average fluctuation in relative power	.57	.10
	Frequency of appeal by all associates	.82	.01
	Frequency of appeal by adults	.68	.05
	Frequency of resistance by adults	−.61	.05
	Frequency of resistance by mothers	−.71	.05
	Frequency of avoidance by adults	−.71	.05
	Average plus evaluation by fathers	−.62	.10
Social behavior			
	Frequency of appeal to all associates	−.61	.05
	Frequency of compliance with all associates	.67	.05
	Frequency of compliance with adults	.71	.05
	Frequency of nurturance of adults	.64	.05
	Frequency of nurturance of children	.67	.05
	Average plus evaluation of fathers	−.62	.10
	Average plus affection for children	−.60	.05
Social interaction			
	Frequency of conflict with children	−1.00	..
	Frequency of conflict with adults	−.57	.10

Other children waxed while adults in general and parents in particular waned as associates of the children. Associate groupings of two or more persons increased. Relative power of the children with associates en masse rose. So did fluctuations from episode to episode in relative power, as the ratio of persons younger than the child to persons older than the child necessarily increased. Appeals to the children went up in frequency while resistance against them and avoidance of them went down. On the child side, appeals to others went down while nurturance of others and compliance with others went up. Conflict decreased. Yet positive evaluation fell on both sides in relations of the children with their fathers; and affection of the children for child associates declined.

The children evidently did not outgrow their control relation with adult associates. In particular, neither dominance by adults nor submission to them by children declined. Some loosening of adult control is suggested by the decrease in resistance by adults. But the basic control relation evidently persisted. And so also did the benevolent side of the dependency relation with adults, inasmuch as grownups nurtured the older children no less than the younger. The older children, though, appealed to adults less, and nurtured them more. It is as if grownups persisted in saying as the children gained in size and strength, "You still need us as much as ever," while the children said, "We need you less" and "You need us more."

These relationships agree generally with sense and one another. They contribute some intelligible measures of increasing social maturity and of compatible change in naturally occurring conditions of the social environment.

Discussion

Biological ecologists have prodigious knowledge of how fish grow in streams, trees in forests, and the like. But there have been virtually no psychological ecologists, with the result that we have known little about the psychological growth of persons in communities. The data here add some knowledge of this kind from the one community of Midwest.

Animal and plant ecologists all have had to adopt standard units of description, such as *grazing ranges, behavior cycles, successions,* and *life cycles*. This need we have tried to meet by choosing as descriptive units behavior settings and episodes. Our first claim for these targets of a psychological ecology would be that description in their terms can take into account both psychological events and conditions under which these same events occur for the reason that each target is a unit of behavior and its context.

Our first claim for the present data, too, would be that they describe developmental change in context or situation as well as in behavior. Procedures in laboratories, testing rooms, and clinics have produced most of our scientific knowledge about psychological growth. These procedures have

taught us much about change in behavior with development, as witness scores of age curves on motor behavior, language behavior, learning behavior, emotional behavior, social behavior. But age curves on change in psychological situation with development are rare indeed. The main reason for this we take to be that the situations in which behavioral growth occurs cannot often be brought into laboratories, testing rooms, or clinics, but generally must instead be left in the field, where they can be studied only by field methods. Some analogues of naturally occurring situations can, of course, be arranged and observed, as rivers are simulated and their laws explored, by experiments that manipulate what they study. This has worked well in psychology. Yet, in the first place, some naturally occurring psychological conditions are like some conditions of weather cycles, ocean currents, and bird migrations in that they cannot be arranged, easily or at all; and, in the second place, such experimental investigation can never solve the purely descriptive, ecological problems which seem to us important in their own right, as they have been in other biological and social sciences.

Also to be considered here is the prospect that psychological field methods can be used beyond descriptive study for inquiry into interdependencies of the sort generally associated with laboratory research. With the present methods, for example, one can explore relations within settings, within episodes, or across the boundaries of settings and episodes. This has had to wait until now on the prior task of straight description with its continuing problems of method. But some such line of investigation via analytical field methods of one kind or another is encouraged in principle by the accomplishments of public health researchers, agronomists, geologists, anthropologists, astronomers, and other descriptive scientists who regularly educe basic statements of interdependence by interrelating naturalistic data.

This leaves implicit in the main idea of psychological ecology, research on behavior and its conditions in differing communities of different cultures. Experiments in nature on consequences for psychological habitat and behavior of differences between communities in size, independence, completeness, population density, rural or urban character, suburban locale, technological advancement, social organization, and various culture dimensions are logical possibilities of such comparative investigation. Whether conceptual and technical requirements of these experiments can be met is an important question which the present methods and data, added to those of earlier field studies in social science, may help to answer.

References

1. Baldwin, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.

2. Barker, R. G., & Wright, H. F. *One boy's day.* New York: Harper, 1951.

3. Barker, R. G., & Wright, H. F. *Midwest and its children.* Evanston: Row, Peterson, 1955.

4. Bates, M. *The nature of natural history.* New York: Scribner, 1950.

5. Kounin, J. S. Intellectual development and rigidity. In R. G. Barker, J. S. Kounin, & H. F. Wright (Eds.), *Child behavior and development.* New York: McGraw-Hill, 1943.

6. Lewin, K. *Principles of topological psychology.* New York: McGraw-Hill, 1936.

7. Lewin, K. *Field theory in social science.* New York: Harper, 1951.

8. Newton, M. R. A study in psychological ecology: the behavior settings in an institution for handicapped children. Unpublished master's thesis, Univer. of Kansas, 1953.

9. Tolman, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century, 1932.

10. Warner, W. L., Meeker, Marchia, & Eells, K. *Social class in America.* Chicago: Science Research Associates, 1949.

11. Wright, H. F., & Barker, R. G. *Methods in psychological ecology.* Lawrence, Kansas: Department of Psychology, Univer. of Kansas, 1950.

12. Wright, H. F., Barker, R. G., Nall, J., & Schoggen, P. H. Toward a psychological ecology of the classroom. *J. educ. Res.,* 1951, 45, 187-200.

CURRENT MONOGRAPHS
of the Society for Research in Child Development, Inc.

Lloyd M. Dunn and Rudolph J. Capobianco
STUDIES OF READING AND ARITHMETIC IN MENTALLY RETARDED BOYS
1956, $3.00

Pauline Beery Mack, Alice Knapper Milsom, and Paul L. Carney
A STUDY OF TWO LEVELS OF BREAD ENRICHMENT IN CHILDREN'S DIETS
1955, $2.00

Douglas M. More
DEVELOPMENTAL CONCORDANCE AND DISCORDANCE DURING PUBERTY AND EARLY ADOLESCENCE
1955, $3.00

Theron Alexander
THE ADULT-CHILD INTERACTION TEST: A PROJECTIVE TEST FOR USE IN RESEARCH
1955, $4.50

Sibylle Escalona, Mary Leitch, and Others
EARLY PHASES OF PERSONALITY DEVELOPMENT: A NON-NORMATIVE STUDY OF INFANT BEHAVIOR
1953, $1.00

Lawrence K. Frank, Ross Harrison, Elisabeth Hellersberg, Karen Machover, and Meta Steiner
PERSONALITY DEVELOPMENT IN ADOLESCENT GIRLS
1953, $4.50

Heinz Werner and Edith Kaplan
THE ACQUISITION OF WORD MEANINGS: A DEVELOPMENTAL STUDY
1952, $1.75

Frank K. Shuttleworth
THE ADOLESCENT PERIOD: A PICTORIAL ATLAS
1951, $2.00
THE ADOLESCENT PERIOD: A GRAPHIC ATLAS
1951, $3.00
Both monographs—$4.00 the set

(A mimeographed list of all monographs available in the series since 1935 will be sent upon request.)

order from
CHILD DEVELOPMENT PUBLICATIONS
Purdue University, Lafayette, Indiana